**U.S. DEPARTMENT OF COMMERCE**
**National Technical Information Service**

PB-259 600

# PROCEEDINGS OF THE 1ST US/USSR SYMPOSIUM ON COMPREHENSIVE ANALYSIS OF THE ENVIRONMENT HELD AT TBILISI, GEORGIA, ON MARCH 25-29, 1974

ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C.

MARCH 1974

**TECHNICAL REPORT DATA**
(Please read Instructions on the reverse before completing)

PB259600

|  |  |  |
|---|---|---|
| 1. REPORT NO.<br>EPA-600/9-75-004 | 2. | 3. RECIPIENT'S ACCESSION NO. |
| 4. TITLE AND SUBTITLE<br>PROCEEDINGS OF THE FIRST US/USSR SYMPOSIUM ON COMPREHENSIVE ANALYSIS OF THE ENVIRONMENT |  | 5. REPORT DATE<br>March 1974<br>6. PERFORMING ORGANIZATION CODE |
| 7. AUTHOR(S)<br>Compiled by L. D. Attaway and R. S. Cortesi |  | 8. PERFORMING ORGANIZATION REPORT NO. |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Office of Research and Development<br>U. S. Environmental Protection Agency<br>Washington, DC 20460 |  | 10. PROGRAM ELEMENT NO.<br>11. CONTRACT/GRANT NO.<br>In-House |
| 12. SPONSORING AGENCY NAME AND ADDRESS<br>Same |  | 13. TYPE OF REPORT AND PERIOD COVERED<br>Symposium<br>14. SPONSORING AGENCY CODE<br>EPA-ORD |

15. SUPPLEMENTARY NOTES
See also Second Joint (EPA-600/9-76-024)

16. ABSTRACT

These proceedings include the 24 papers presented at the first Symposium. The papers were originally presented in English and Russian at Tbilisi, Georgia, USSR, between March 25 and March 29 1974. Their publication herein is part of the protocol agreement signed on March 29, 1974, for simultaneous and independent publication in both countries after coordination of the manuscripts. Papers presented were grouped into the following categories: Theme-Analysis to Support Environmental Management; Comprehensive Modelling; Maximum Permissible Loading; Economic Criteria for Establishing Environmental Quality Goals; Implementation Strategies for Achieving and Maintaining Environmental Standards; Impact of Pollution and Environmental Conditions on Non-Human Species; and Pollution Impacts on Climate.

17. KEY WORDS AND DOCUMENT ANALYSIS

| a. DESCRIPTORS | b. IDENTIFIERS/OPEN ENDED TERMS | c. COSATI Field/Group |
|---|---|---|
| Bioclimatology<br>Biochemical oxygen demand<br>Environmental surveys<br>Mathematical models<br>Ecology<br>Pollution | USSR<br>Biosphere<br>Maximum permissible load | 04B<br>06F<br>12A |

|  |  |  |
|---|---|---|
| 18. DISTRIBUTION STATEMENT<br>Unlimited | 19. SECURITY CLASS (This Report)<br>Unclassified | 21. NO. OF PAGES<br>191 |
|  | 20. SECURITY CLASS (This page)<br>Unclassified | 22. PRICE<br>$7.50-3.00 |

EPA Form 2220-1 (9-73)

[illegible]

[illegible]

PB 259 600

# PROCEEDINGS OF THE FIRST US/USSR SYMPOSIUM ON COMPREHENSIVE ANALYSIS OF THE ENVIRONMENT

JOINT SOVIET-AMERICAN SYMPOSIUM ON THE COMPREHENSIVE ANALYSIS OF THE ENVIRONMENT
СОВМЕСТНЫЙ СОВЕТСКО-АМЕРИКАНСКИЙ СИМПОЗИУМ ПО ВСЕСТОРОННЕМУ АНАЛИЗУ ОКРУЖАЮЩЕЙ СРЕДЫ

MARCH 25-29, 1974

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

EPA-600/9-75-004

# PROCEEDINGS OF THE FIRST US/USSR SYMPOSIUM ON COMPREHENSIVE ANALYSIS OF THE ENVIRONMENT

# Preface

These proceedings include the 24 papers presented at the first U.S./USSR Symposium on Comprehensive Analysis of the Environment. The papers were originally presented in English and Russian at Tbilisi, Georgia, USSR, between March 25 and March 29, 1974. Their publication herein is part of the protocol agreement signed on March 29, 1974, for simultaneous and independent publication in both countries after coordination of the manuscripts.

# Acknowledgement

The success of the first symposium on Comprehensive Analysis of the Environment resulted from the efforts of many persons who planned the agenda, arranged facilities, coordinated papers and their translation, and attended to the myriad of details which occurred through the symposium. The productivity of the 1974 symposium and its fruitful exchange of ideas were a vivid realization of the objectives in the Agreement on Cooperation in the Field of Environmental Protection, signed in Moscow on May 23, 1972.

Achievement of these plans is a special credit to the Soviet delegation under their leader, Professor Yu. A. Izrael, First Deputy Chief of the Main Administration of the Hydrometeorological Service, Council of Ministers of the USSR. Since the Tbilisi symposium, now Academician Izrael has succeeded Academician E. K. Fedorov as Chairman of the Soviet Side of the Joint Committee on Cooperation in the Field of Environmental Protection. The warm appreciation for all his help has been duly noted by the head of the American delegation, Dr. Stanley M. Greenfield, Assistant Administrator for Research and Development, U.S. Environmental Protection Agency. The American contributions were compiled by Dr. Leland D. Attaway, Deputy Assistant Administrator for Program Integration, U.S. Environmental Protection Agency. Completion of these proceedings has been directed by Roger S. Cortesi, Director of the Washington Environmental Research Center, U.S. Environmental Protection Agency, who succeeded Dr. Attaway as Chairman of this U.S. Working Group Project.

# Contents

*Page*

## I. THEME-ANALYSIS TO SUPPORT ENVIRONMENTAL MANAGEMENT

Comprehensive Environmental Management, an Overview—S. M. Greenfield ..................... 1

## II. COMPREHENSIVE MODELLING

Systems Analysis and Simulated Mathematical Modelling as a Methodological Basis for the Standardization of Anthropogenic Approach—Yu. A. Izrael, Yu. A. Anokhin ................ 5
A Comprehensive Environmental Analysis of the Upper Potomac Estuary Eutrophication Control Requirements—N. A. Jaworski ..................... 15
The Ecological System Stability Concept—V.D. Fedorov ........ 27
Development of Maximum Permissible Environmental Loading-Air—B. J. Steigerwald ..................... 35

## III. MAXIMUM PERMISSIBLE LOADING

Estimating the Maximum Permissible Intensity of the Complex Effect of Chemical Factors in the Production, Municipal, and Domestic Environment on Man—N. F. Izmerov .................... 45
The Concept of the Threshold Nature of Reaction of Living Systems to External Actions and its Consequences in the Problem of Protecting the Biosphere against Chemicals—I. V. Sanotskiy ..................... 49
Questions of Extrapolating Data in Evaluating the Mutagenic Action of Environmental Factors on Man—N. P. Bochkov .... 55
The Question of Genetic Danger from Environmental Pollutants—L. M. Filippova ..................... 59
Hygienic Bases for Protecting the Environment—G. I. Sidorenko, Ye. I. Korenevskaya, G. N. Krasovskiy .................... 63
Hygienic Normalization under Scientific and Technical Progress Conditions—A. P. Shitskova, M. I. Gusev, S. M. Pavlenko, I. L. Karagodina ..................... 67
Maximum Permissible Human Stress—V. Newill .................... 71
Maximum Permissible Human Stress with Multiple Pathways: Ionization Radiation, Asbestos and Lead—J. H. Knelson ........ 81
Complex Analysis of the Environment—Approaches to a Determination of the Permissible Loads on the Natural Environment and to a Justification for Monitoring—Yu. A. Izrael .... 89
Defining "No Significant Deterioration" of Air Quality—R. A. Luken ..................... 95

## IV. ECONOMIC CRITERIA FOR ESTABLISHING ENVIRONMENTAL QUALITY GOALS

Determining Acceptable Levels of Health and Environmental Damages—F. H. Abel ..................... 107

**Preceding page blank**

Page

V. IMPLEMENTATION STRATEGIES FOR ACHIEVING AND MAINTAINING ENVIRONMENTAL STANDARDS
Methods for Abatement of Water Pollution—R. B. Schaffer .... 117
Methods for Control of Air Pollution—K. E. Yeager ............ 125
Development of Maximum Permissible Environmental Loading (MPEL)-Water—M. A. Pisano .......................................... 139

VI. IMPACT OF POLLUTION AND ENVIRONMENTAL CONDITIONS ON NON-HUMAN SPECIES
The Effect of Polluting a Biogeocoenosis with Strontium 90 on a Population of Mammals and the Zoocoenosis—V. Ye. Sokolov, A. I. Ilyenko, D. A. Krivolutskiy ............................ 149
Effects of Pollution on Species and Populations of Fish and Birds—H. E. Johnson .............................................................. 159
The Influence of Environmental Factors on Population Dynamics (Mathematical Models)—M. Ya. Antonovskiy .................... 171

VII. POLLUTION IMPACTS ON CLIMATE
Global Climate and Human Activity—M. I. Budyko, I. L. Karol 179

VIII. SUMMARY AND RECOMMENDATIONS
Concluding Remarks for the U.S./USSR Symposium on Comprehensive Analysis of the Environment—L. D. Attaway ........ 187

# Attendees

## AMERICAN PARTICIPANTS

| | |
|---|---|
| Dr. Stanley M. Greenfield | Head of Delegation<br>Assistant Administrator for Research and Development, USEPA |
| Dr. Leland D. Attaway | Co-chairman, Comprehensive Analysis of the Environment Project<br>Deputy Assistant Administrator for Program Integration, USEPA |
| Dr. Frederick H. Abel | Chief, Economic Analysis Branch, Washington Environmental Research Center, USEPA |
| Dr. Norbert A. Jaworski | Director, Pacific Northwest Environmental Research Laboratory, USEPA |
| Dr. Howard E. Johnson | Private Citizen, Associate Professor, Department of Fisheries and Wildlife Michigan State University |
| Dr. John H. Knelson | Acting Director, Human Studies Laboratory, USEPA |
| Dr. Ralph A. Luken | Private Citizen, Consultant on Environmental Economics |
| Dr. Vaun Newill | Special Assistant to the Administrator, USEPA |
| Dr. Mark A. Pisano | Director, Water Planning Division, USEPA |
| Mr. Robert B. Schaffer | Acting Director, Permit Assistance and Evaluation Division, USEPA |
| Dr. Louis J. Schoen | Special Assistant for International Research and Development, USEPA |
| Dr. Bernard J. Steigerwald | Director, Office of Air Quality, USEPA |
| Mr. Kurt E. Yeager | Chemical Engineer, Office of Research and Development, USEPA |

## SOVIET PARTICIPANTS

| | |
|---|---|
| Prof. Yu. A. Izrael | Co-chairman, Comprehensive Analysis of the Environment Project<br>First Deputy Chief of Main Administration of Hydrometeorological Service, Moscow |
| Dr. Yu. A. Anokhin | Institute of Applied Geophysics, Moscow |
| Prof. M. Ya. Antonovsky | Central Institute of Economic Mathematics, Moscow |

| | |
|---|---|
| Prof. M. E. Berljand | Main Geophysical Observatory, Leningrad |
| Prof. M. I. Budyko | Main Geophysical Observatory, Leningrad |
| Prof. V. D. Fedorov | Moscow State University, Moscow |
| Dr. L. M. Filippova | Institute of Applied Geophysics, Moscow |
| Prof. M. I. Gusev | Institute of Hygiene, Moscow |
| Prof. G. F. Hilmi | Institute of Applied Geophysics, Moscow |
| Prof. A. I. Ilyenko | Institute of Evolutional Morphology and Ecology of Animals, Moscow |
| Dr. N. F. Izmerov | Institute of Labour Hygiene and Professional Diseases, Moscow |
| Dr. V. S. Jurkov | Institute of Medical Genetics, Moscow |
| Prof. I. L. Karol | Main Geophysical Observatory, Leningrad |
| Dr. Yu. E. Kazakov | Joint Soviet-American Committee, the Soviet Part of the Project |
| Dr. G. N. Krasovsky | Institute of General and Communal Hygiene, Moscow |
| Dr. B. A. Kuvshinnikov | Joint Soviet-American Committee, the Soviet Part of the Project |
| Prof. M. J. Lemeshev | Central Institute of Economic Mathematics, Moscow |
| Dr. V. P. Lominadze | Director of Transcaucasian Hydrometeorological Institute, Tbilisi |
| Prof. A. M. Molchanov | Computer Centre, Puschino-na-Oke |
| Dr. I. M. Nazarov | Institute of Applied Geophysics, Moscow |
| Dr. V. G. Novozhilov | Assistant Chief of Main Administration of Hydrometeorological Service, Moscow |
| Prof. I. V. Sanotsky | Institute of Labour Hygiene and Professional Diseases, Moscow |
| Acad. S. S. Shvarts | Institute of General Biogeocenology and Ecology, Sverdlovsk |

# Introduction

The Joint Committee on Cooperation in the Field of Environmental Protection was established by the signing of an agreement in Moscow, May 23, 1972. This was followed by the Second Session of the U.S./USSR Joint Committee in the Field of Environmental Protection (Washington, D.C., November 16, 1973).

From these organizational meetings, eleven areas were specified for cooperative investigation and exchanges. Within these broad areas a number of projects have been identified. These proceedings are based on one of these projects: Project VII–2.1, Comprehensive Analysis of the Environment, within Area VII—Biological and Genetic Ecects of Environmental Pollution. In accordance with provisions of the Memorandum of the Second Session, the symposium at Tbilisi was scheduled to bring together experts from both countries for presentation and discussions of papers in this area. This area is unique in the comprehensive perspective of its approach to environmental problems; consequently the contents or titles of some of the papers could suggest that they are more germane to other program areas.

The proceedings are arranged to present papers in topical clusters, and do not necessarily reflect the original sequence of presentations. Papers from each delegation are intermingled to emphasize the contrast in approaches, and suggest the lively discussions which they engendered.

Primary emphasis was given to the determination of scientifically permissible loading on the environment. This emphasis considered the comprehensive analysis of the pollution sources and receptors as well as approaches to the development of criteria for scientifically permissible stress to the environmental and identification of critical links in this development.

The Tbilisi symposium identified the following problems which will require attention:

(1) More complete accounting of pollution sources and receptors, types of impacts and biological reactions, as well as peculiarities of the area or region where such impact occurs;

(2) Investigation of dose/response relationships in toxicological, genetic, and other experiments;

(3) Scientific substantiation of environmental monitoring techniques with due consideration of critical links in analyzed systems;

(4) Studies of the fate and effects of multimedia pollutants, supplemented by development of the required standards;

(5) Assessment of the needs for separate standards limiting long- and short-term dosages of selected environmental pollutants;

(6) Methods for evaluaing the response of ecosysems to varying environmental conditions;

(7) Development of criteria for scientifically permissible environmental loading and identification of critical links;

(8) Development of practical quantitative methods which will permit comprehensive environmental analysis to be applied effectively to environmental problems;

(9) Development of methodology for assessment of possible economic damage resulting from man's impact on the environment; and

(10) Feasibility of a joint U.S.–USSR monitoring system.

The next meeting on the Comprehensive Analysis of the Environment, or Symposium II, is scheduled for the Fall of 1975, and will have as its theme:

Maximum Acceptable Environmental Stress on Organisms, Populations, Ecosystems and the Biosphere as a Whole.

# Comprehensive Environmental Management—An Overview

Stanley M. Greenfield *

From a regulatory sense, it is possible to separately consider individual pollutants emerging from single sources into a single medium. Experience has shown, however, that to derive a strategy that will ultimately result in environmental quality, one must comprehensively consider the entire complex interacting environmental system within the context of operative socio-economic constraints. Such consideration must take into account the distribution of pollutant sources throughout the spatial region in question, and their intensity variation with time. What then must be derived is a mechanism for specifying the required temporal and spatial variation in effluent so as to produce the desired environmental quality. Implicit in such a determination is the recognition that ultimately there must be a balance in the cost of achieving this level of quality and the benefit that accrues to society in its achievement. All of human activity, in one way or another stresses the environment. The question is, how much stress is acceptable, and how is such a limit achieved and maintained? Obviously, within limits, a society could choose the degree of environmental management it desires, the limits on the choice of degree being when it permits those levels of stress to occur which produce irreversible, deleterious impacts on the society itself. The basic question at issue, however, is how does the society properly make such a choice? And, within the context of that question, what information is required to make such a decision? What trade-offs are to be considered? What are the potential methods of implementation of any given management strategy? What analytical tools are available and/or required? What is the proper relationship between short-term goals to accomplish an immediate end, and long-term goals to find a lasting solution?

First on any list of information required for such a decision process must be some determination of the maximum permissible species stress. This concern is conveniently divided into two basic areas, first the effects associated with human health, and second, those associated with the other segments of the ecological system.

In the case of health effects, we immediately face the dilemma of social risk and its acceptable level. Are we to consider the economic questions associated with protecting populations, and are these to enter into any determination of maximum permissible stress? The first inclination is to dismiss economic and social risk as an inappropriate consideration within any discussion of human health. This is particularly easy to do when one is faced with a health effect threshold clearly defined and applicable to the population as a whole. In general, however, this simple case is seldom realized. What we are normally faced with is either a dose-response curve with considerable uncertainty near zero dose, or a threshold that varies with different subgroups of the population (e.g., cardiovascular sufferers, etc.). The dilemma is clear in both cases. It is possible or desirable to provide a level of maximum permissible risk that is not protective of the entire population? Conversely, is it desirable or necessary to protect down to the last noticeable effect, when to do so may require literally the forbidding of the flow of any quantity of a particular pollutant into the environment over and above the natural background?

In addition to the relatively simple case of the single pollutant as a unique stressor in a unique pathway, we must consider the implication of the synergistic and nonsynergistic effects of multiple stressors, and the equally complex multiple pathway single pollutant case. In the first case, one is possibly implying a lowest common denominator maximum permissible stress that depends on what other pollutants are present in the environment. In the latter case, such as lead and its many pathways to man, we are faced with a maximum permissible stress depending on the mechanism of uptake and some consideration of the relative intensity of various intake models available.

The data that underlie the determination of the maximum permissible stress for a given pollutant are not particularly impressive. They become even less impressive when one considers multiple pathway and/or multiple pollutants. Epidemiologic experiments are difficult to carry out and when stratified to reveal the single or specific multiple pollutant effect, quickly reduce even a large body of data to an insufficient sample. Animal toxicological experiments are difficult to carry out successfully for the long

---

*Assistant Administrator for Research and Development, U.S. Environmental Protection Agency periods of time required, and suffer in their translation to man. Yet, the ability to state the maximum permissible stress is the cornerstone to being able to describe an adequate control plan, and to provide the input options for an adequate environmental management system. In addition, adequate ability to quantify the stressor-effect relationship is mandatory if we are to ultimately describe the social consequences of various levels of pollutant concentrations.

Of particular interest here is the question of nondegradation, or the decision to not permit a region that is currently below the maximum permissible stress to degrade so as to reach that limit. This is an extremely difficult question to deal with from both a philosophical and social standpoint. It starts with the assumption that we have established a maximum permissible stress. Philosophically or pragmatically, do we now allow all areas to be polluted up to that point or are we determined to hold the environmental quality of relatively unimpacted areas? The former implies that the maximum permissible stress represents a level to which we will allow all levels to rise. The cost of nondegradation must be very carefully weighed against the social benefits that accrue.

All of the issues raised in describing the process of obtaining a maximum permissible environmental stress for health effects are applicable to the case of ecological effects as well: threshhold versus nonthreshold pollutants, benefits of given levels of maximum permissible stress, defense of especially sensitive subpopulations, simple pollutants, multiple stresses, both synergistic and nonsynergistic, and multiple pathways. In general, however, the ecological effects require an even greater consideration of the social benefits that accrue to a given level of permitted stress or control. Obviously, we are dealing with a larger set of intangibles as opposed to the relatively more easily defined area of health protection. The result is that a society which can understand and accept costs attached to protecting human health has a great deal of trouble justifying large costs to protect ecological systems. The interrelated character of our environmental system is little understood outside of the scientific disciplines. The idea that damaged species in an ecological system can produce irreversible effects in the entire system is difficult to grasp and hence, demands an overt attempt to trace the relationship through to a clearly defined societal cost and benefit. The one area where this is possibly not true in the consideration of ecological effects is that involving impacts on the food chain that can more easily be visualized by the layman as having a definable effect on man. These effects, of course, are even more difficult to deal with from a technical sense and involve an in-depth consideration of the complex multiple pathways, multiple stressor interactions. However, the ability to demonstrate that man is destroying his food web has a profound effect that leads to a recognition in society of the benefits that accrue and a willingness to accept the high "cost" of prevention.

To this point, we have been discussing the pollutants, the pathways, and the maximum stress allowable through these pathways so as to protect various species. The information obtained must provide a sense of the levels of control necessary in these pathways to provide the required protection, and a sense of the population at risk as a function of the applied controls. In essence, this represents the first step in a decision making process leading to effective management of the environment. The decision maker, however, must have the ability to address the control options available to him and the cost, both social and monetary, that apply to the various levels.

There are two basic control options available—technological, and process or operations management. In the technological case, we are referring to specific waste stream controls that can be applied to specific sources. Typically, these are stack gas cleaners, biological treatment plants, etc., hardware solutions for clearly defined waste streams. The process or operational option refers to changes in the way the pollutant is produced so as to limit the effluent. In the case of a specific plant, this could involve a change in an internal chemical process (e.g., substitution for the mercury cell in a chlor-alkali plant). In the case of a widespread source, such as transportation, it could involve a management process that results in a change in use pattern. Sources considered can be either stationary or mobile, located at a point as in the case of a power plant or widespread as in the case of agriculture. Pollutants can be those that are emitted in final deleterious form (e.g., radioactive gases, carbon monoxide, etc.) or take form after emission (e.g., photochemical smog, sulfates, etc.).

The acceptability of a given control option is dependent on both its dollar and social cost. A typical example of a dollar and social cost is the combined cost of restricting the use of automobiles and the social acceptability or unacceptability of such constraints. These must depend on the level of stress limitation desired. Yet, this is not an explicit factor in the sense that to some undetermined degree the maximum permitted stress is dependent on the cost that society must pay. The decision maker in managing the environment requires, therefore, a determination of the economic costs of various options as a function of desired control level as well as some estimate of the variation in the institutional impacts that result. This then provides him with the ability to take the second step in the decision making process, the information that will permit option trade-offs to be examined.

The final step in this process of developing an environmental management plan involves the development of an operational strategy within which the various trade-offs can be examined. In essence, these sets of strategies represent control plans involving the effluent quality and its variation with time for all sources distributed over the entire region. The strategy, therefore, represents a plan for limiting the stress to the maximum permissible by limiting the maximum permissible loading, and distributing this over the region by assigned source effluent limits. Obviously, adequate knowledge of control costs, stress impact, and source distribution in time and space will permit the construction of management strategies that can be optimized either singly or in combination around various conditions (i.e., cost, maximum protection, etc.). The independent variable in all cases is the date by which a given level of protection will be achieved.

Strategies for managing the environment can be subdivided into two basic classes, the first being the case described previously of preventing an area that is currently clean from ever reaching the maximum allowed stress and the second being the implementation of strategies that force a currently polluted area to recover a quality environment at least equal to the maximum permissible stress.

In the first case, the point at issue concerns the price that society is willing to pay to keep a clean area clean. Within this issue is the requirement to define what we mean by clean. We presumably have already established a level of maximum permissible stress. Further, we have already indicated that level in all probability is not an absolute protection for all affected species. Does this imply that nondegradation means zero increase above present levels? Or conversely, does it imply an allowable increase that will not permit the stress to exceed a stated fraction of that which is the maximum permissible? The first case will demand a cost which will probably make it impossible for any additional development to occur. The second might permit some limited development, but at a stated cost that may or may not be unacceptably too high. It is clear that with the approval of society, the description of the acceptable nondegradation level can be used as a mechanism for protecting designated areas from development of a certain character, and restrict development to specified areas. It becomes, in effect a deliberate tool for land use management.

It is certainly attractive to consider the ultimate protection of pristine areas and the demand for an absolute nondegradation. It must be pointed out however, that too restrictive a demand can be counterproductive to the environmental needs of society. Imagine, for example, a nondegradation level that is so restrictive as to permit development to occur only in the already polluted and degraded areas. It is clear that such an approach would condemn the highly developed, most populated areas to ever increasing development and stress. Once again, it is necessary to carefully balance the costs and benefits to society of the management actions to be taken in the sense that any management tool is a two-edged sword that badly used can be more damaging than inaction. Used intelligently, the nondegradation issue can help control and alleviate the environmental stress in the more polluted areas, can protect the pristine areas, and can provide for an orderly growth in the designated regions.

In the second case, the case of environmental recovery, we are attempting to work with an area that is normally heavily populated, heavily developed, and most probably of strong economic importance. The problem the environmental manager faces is how to design a program of distributed control for the region that takes into account the temporal and spatial distribution of sources and further, that this program of control provides the region with an environmental quality that is within the maximum permissible stress without destroying its socio-economic viability. In addition, a decision must be made as to the desired date of achievement of this level of control. The control plan evolved and the attainment schedule chosen are obviously related when considered in the context of cost-benefit balancing. This is true both in the sense of the stipulated maximum permitted stress as indicated previously, and in the distribution of control level and its designated time of application. The ability to do so, however, implies a capability to assess the social risks of a given stress level and its period of maintenance, and to accurately assign a stress level contribution to each source. Under the influence of total cost as related to benefits, one might stipulate a decreasing maximum permissible stress with increasing time, paced to match into the achievement ultimately of a more effective control capability with a smaller societal cost. Obviously, this must also be paced with the cumulative social disbenefits that may increase with the delay in the achievement of the ultimate maximum permissible stress. What we are searching for then is the ability to minimize the sum of direct abatement costs (whether actual economic costs for control equipment or social costs due to operational changes, such as reduced automobile use), and the social costs of delaying the achievement of the most desirable maximum permissible stress.

One important note of caution must be sounded in the above approach. Unless considered directly, and included in the management plan, one may find that this approach contains no incentive for ever acquiring the techniques for achieving the most desired maximum permissible stress. The approach utilized to date, particularly in the United States, has been directed primarily towards setting maximum permissible stress on the more restrictive grounds of environmental effects as determined, and the achievement date as a legislated mandate. What cost-benefit analysis is done, is normally done after the fact of setting the values of these parameters. In defense of this system: First, it does not require knowledge, as described above, that is not currently available to the extent needed; hence, it does not delay the regulatory process. Second, it does force the development of techniques to achieve the desired levels. The undesirable feature of this system is that it may be counterproductive in its attempt to protect society and at the very least, it is an abrasive system that puts a premium and a relatively low penalty on noncompliance by individual sources.

The inadequacy of the current regulatory system is rather widely recognized, and gains in recognition with each further thrust towards comprehensive environmental management. This is not to say that a great deal has not been accomplished with the current approach. This is particularly true in the sense of bringing about a problem awareness, underscoring the government's resolve to achieve an improvement in the environmental quality, and setting the research program on the proper course. As the problems become more complex, however, as the societal problems become more massive and evident, the realization of the need for a more analytical comprehensive approach grows. As complexities grow, so grow the questions that must be addressed. Should there be ambient standards, effluent standards, or both? Should health effects carry the same weight as other ecological effects, can and should cost-benefit analysis be used to set the standards, the date of achievement, or both? Should standards be expressed as a maximum allowable concentration over a specified period, and if so with what frequency? Are multiple pollutant standards required, and how will goal achievement be determined?

Along with the realization of the inadequacy of the simplistic direct enforcement approach to deal with complex environmental management problems goes an awareness of the accompanying inadequacy of the needed analytical tools. In order to specify a control plan and an achievement schedule which truly minimizes socio-economic costs (e.g., minimizes the societal cost-benefit ratio), it must be possible to relate that plan to species stress and, therefore, to ambient quality. The preferred method of doing this would be through the use of stochastic physical-bio-chemical models for both water and air which adequately represent the probability aspects of local through regional transport and bio-chemical processes. Although much attention has been given and is being given to the development of such models and the required input data, neither is adequate at this time. As a result, interim methodologies and approaches are utilized, which fall far short of being able to adequately address the questions raised. It is certainly clear that a substantial effort is needed in the near term to address the highly impacted areas. A major concern is felt, however, lest actions taken now utilizing gross data and gross techniques preclude more correct solutions being implementable at a later date.

A final point of concern to the environmental manager is involved in the achievement of a regional long-term plan of environmental control that meshes well with the pressing short-term needs. Short-term programs can be expected to have less stringent goals than the long-term ultimates, and rely more on immediately available controls than on longer range technological and operational fixes. Environmental pollution, however, is a symptom of the intensity and scale with which human activities are pursued in the region. Each region has a natural "carrrying capacity" for these activities, which can be altered significantly only over a long period of time. For example, the carrying capacity can be altered by technology (e.g., mass-transit to off-set the use of automobiles; or pollution control technology) by urban design and land-use planning, and by resource conservation measures, such as improved building insulation and design. Long-terd environmental programs in effect are increasing this regional carrying capacity by altering the basic nature of the region and its socio-economic activities. On the other hand, much of a short-term program (e.g., transportation controls) is simply reducing the intensity or scale of human activities, within a fixed carrying capacity. Ultimately, every region must live within, and plan for, a finite carrying capacity. In particular, this means that social and economic goals must be matched to the natural characteristics of the region involved and the interaction and impact of one region on another.

In summary, then, the ability to effectively manage the environment comprehensively involves many factors. First it must be managed over a region large enough to permit the effective control of the operative pollutants. Second, it is necessary to determine the maximum permissible environmental stress to protect the effective health of the population and the underlying ecological system In this connection, questions of social risks should be addressed, and the concept of total benefit considered. Third, the cost of control plans must be adequately known. Fourth, a maximum loading plan must be established which permits a consideration of source distribution and intensity, as well as the social costs of control. Fifth, a long-term plan compatible with short-term needs must be evolved which, if necessary, can allow the region to reach the limits of its natural carrying capacity. These parameters, adequately treated, will permit a management plan to be developed which addresses the ability to comprehensively protect against the environmental impact balanced with the necessity for society to remain viable.

This symposium obviously will not provide answers to all of the questions raised, but will allow both countries to share the knowledge we have both acquired and explore what has yet to be done—it is a beginning.

# Systems Analysis and Simulated Mathematical Modeling as a Methodological Basis for the Standardization of Anthropogenic Pollution of the Environment: A Regional Approach

Yu. A. Anokhin, Yu. A. Izrael

1. One of the serious negative consequences of man's activities (primarily economic), as is known, is the pollution of the environment, the scales of which grow rapidly, but which, nevertheless, ought not be considered a necessary satellite of development [1]. Anthropogenic pollution of the environment occurred in the course of historical development, and is a unique (albeit far from the best) type of "exploitation" of natural resources—the air reservoir, natural waters, and soils.

Scientists and specialists, while acknowledging the social origin of the problem of environmental pollution ought not, however, limit themselves simply to calling for new ethical standards to apply to man's relationships with nature. Our task is to develop effective, scientifically valid, measures for dealing with pollution, and preventing the occurrence of possible crisis phenomena. By "pollution" let us understand that we mean the input into the environment of toxic substances that can be harmful to people's health, as well as of any other substances that could have undesirable effects on biogeocenosis and on materials, and this includes excess heat, noise, and radiation in the environment. In truth, impurities are substances (and radiation fields) found where they are not supposed to exist (this is a modified version of an old chemical rule). It would only be natural to include here as well anthropogenic changes in landscapes deemed unfavorable, not only from an economic, but from an esthetic point of view as well.

The possible effects mentioned here are of great social, cultural, and economic importance, and all must be taken into consideration when setting pollution standards, either in terms of the conditions the role of the "leading" object must be in so far as man's health is concerned, the effect it will have on predetermined natural systems (biogeocenosis), on certain forms of flaura and fauna, or even on certain materials.

Moreover, and this is of utmost importance, an understanding such as this of the pollution problem will make it possible to deal with it as one of the facets of the wrong, irrational use of natural resources, so that the approach to solving the problem follows as the antithesis: find ways to make intelligent, "optimum," use of natural resources.

2. A characteristic property of the "optimization" problem under consideration is its internal contradiction, its "multidimensionality," so to speak, something a goal-oriented function takes on when an attempt is made to assign the corresponding extremal problem.

The overall goal, for example, is that of rationalizing the interaction between industry and the environment, and comprises the following sub-goals, the latter in many respects contradictory: maximization of physical production and of the productivity of labor, and minimization of waste products and rejected products. While maximum physical production and consumption could be considered the optimum so far as satisfying human demands is concerned, this approach is not optimum, however, from the standpoint of intelligent use of nature (nor, incidentally, from the standpoint of widely understood social requirements). Natural resources are susceptible to multiple use, so the unresolved problem is "how to make a quantitative comparison of the different social requirements because of the difficulty in comparing the need for clean air, and the esthetic enjoyment of the beauty of nature, with the needs for electric power, metal, and fish, and because of the difficulty, in turn, of comparing the latter each with the other" [2].

It is obvious, therefore, that any attempt to resolve the problem of how to use nature more efficiently "in one fell swoop," by resort to optimization procedures of the mathematical programming type, let us say, is an illusion. An informal understanding of the concepts of "optimum" and "optimization" must be reached by the use of intelligent, "meta-model" considerations, and this approach has a decisive effect on methodology selection.

3. The development of the correct methodology for standardization of anthropogenic pollution is the

most important condition, and means, for dealing successfully with pollution. Systems analysis, and simulated mathematical modeling, as applicable to relatively closed territorial-economic systems, can, in our view, become the instrumentality of the methodology for standardization on a regional scale.

The very fact that environmental pollution is most intensive within the limits of existing territorial-economic regions is why regional introduction of standardization of anthropogenic pollution in accordance with accepted criteria for rational use of nature is so important. It seems clear that the only policy that makes any sense at this time when developing these criteria is that of a differentiated approach to our country's various natural systems (to the biogeocenoses, for example). Some natural communities should be completely cut off from any interference by man, and should be preserved in their natural form. There are others, however, that should be adapted to human needs more decisively and intensively than is the case at the present time. It is possible that there are intermediate types of communities, and here we can permit ourselves to experiment, bearing in mind at all times, however, the possibility of unforeseen, negative, consequences. A policy such as this is optimum in the sense that it guarantees "maximum freedom of choice for our descendants" [3].

4. We shall, in this report, deal first (paras. 5 and 6) with some of the general properties of the systems in the class under consideration, regardless of concrete realization. We then (paras. 7–9) will take up in some detail certain methodological questions concerned with the standardization of pollution on a regional scale, and will illustrate these considerations by the example of the Lake Baykal region (paras. 10–23).

5. As is known, the general properties of such systems, because they are "organized complexity" systems [4], are given theoretical consideration in general systems theory. The most important of these properties are openness (interaction with the environment), the hierarchical nature of the structure, temporal changeability, and the attribute of control inherent in these systems [5–8]. Of significance to us is the fact that these general properties have made it necessary to develop a new methodology for studying complex systems, the basis of which is systems analysis and mathematical modeling, with its general directionality that from analysis (understood to mean the breaking down of the whole into all of its smaller parts and studying each of the latter separately) to synthesis, and the elements of inadequacy and language polymorphism inherent in this method [8, 9].

6. This approach is neither a narrow, special discipline, nor some sort of mystery. It is, rather, a multidiscipline, quantitative method for studying complex systems when knowledge of them is incomplete, the purpose of which is to predict the behavior of the system under changing conditions, and, in a certain sense, to gain optimum control of the system. System behavior is determined by its structure, and by the interrelationships of its elements. The following properties of systems of this class are intrinsic [10]:

(1) they have a great many variables (of the order of several hundred, or more);

(2) the variables often have feedbacks;

(3) the system often has time lags, nonlinearities, and even discontinuities.

Equations describing a system such as this usually are without analytical solution, and this is why it becomes necessary to turn to experimental investigations with mathematical models and resort to an electronic computer. As is known, the Forrester school has gathered a great deal of experience in this field [11–13]. Forrester [11] has made a clear analogy between controlling a complex system and a unique method of driving an automobile, one in which the driver is blindfolded and follows instruction given him by an instructor looking out the rear window (that is, on the basis of information on the road already traveled). Added to this is the fact that the instructions, as given, contain time lags and errors. Given conditions such as these, it is difficult to overestimate the importance of an approach that is based on systems analysis and simulated mathematical modeling, primarily because this approach makes it possible to *predict* possible system states. This, in turn, serves to make system control more effective, and behavior of the controlled system more regular. This approach can, in its general outlines, be said to comprise the following main stages [11, 14]:

(1) the concrete question in need of an answer is formulated. This is basic. Successful formulation of the question is possible only if the situation, the problem, and the purpose of the research, are clearly understood;

(2) a verbal (verbal-causal) model, that is, the fundamental internal and external relationships, is formulated;

(3) the verbal model is graphed, and then is reflected as a mathematical model;

(4) system behavior is computer simulated, and the results are compared with available data on similar real processes;

(5) iterativeness of the process is determined. Return to stage (2) if the representation in the model of real processes is unsatisfactory.

We can obtain the required degree of adequacy of the model to the real system by repeatedly "fitting" the model as we perfect our knowledge to the system's arrangement. And the statement of the problem [(1), above] can be modified as a result of such computer "experimentation."

7. Let us now turn to the methodological questions of pollution standardization. There can be no question of the fact that any program for dealing with environmental pollution must include a complex of sanitary-hygienic, biological, recreational-esthetic requirements, etc., for environmental quality. Nor is there any question of the fact that realization of a program such as this has to depend on social-economic, cultural, and other factors, as well as on the "level of lack of quality" of the environment that in fact exists. The present need for an optimum solution of complicated problems concerned with pollution standardization therefore should be obvious.

Let us note that existing economic and mathematical methods of optimizing the use of production resources cannot be applied directly to solving problems concerned with optimizing the use of natural resources, and this is particularly true when it comes to optimum pollution standardization, because these methods fail to take into account the qualitative specifics of natural resources, which are production factors, as well as the "bodies and forces of nature" [15], "the behavior" of which is determined by natural laws that we know nothing about all too often, unfortunately.

8. Nevertheless, the economic approach, despite all its very serious and well-known limitations, must and can be a powerful lever to use in solving the problems of how to make optimum use of nature. This approach is based on the concept of the reproduction of natural resources as one of the main spheres of physical production [16]. At the same time, the central economic problem is one of a correct *economic* assessment of natural resources. The concept now used by Soviet economists to make this assessment is that of "closing costs," so-called [17, 18].

9. One of the general laws of the behavior of environmental pollutants is that they tend to disperse. Moreover, substances that are reciprocally convertible, particularly when the resultant substance is more toxic than the original one, are of the utmost importance. It is only natural, therefore, that the general principle for the standardization and monitoring of pollutants should be that of localization; that is, the content of pollutants in the environment, and particularly their source, must be standardized and monitored. What follows from this almost self-evident assertion is the practically important conclusion that the presently universally accepted practice of simply standardizing and monitoring the threshold limit value (TLV) is not effective enough. We must clearly realize that there are two closely interrelated aspects of the pollution standardization and monitoring problem. First of all, there is the need to establish the corresponding criteria, or standards, for environmental quality (TLV types, or values similar to them). Absolute conformity of environmental quality with the established standard should be the law, and should not be violated. Realistically, however, the attainability of this state of affairs very definitely depends on standardization and monitoring the very sources of pollutants. This is why the standardization and monitoring of environmental quality on the one hand, and of the sources of pollutants on the other, are mutually complementary.

10. We shall, in this report, for purposes of greater concretization, limit ourselves to the consideration of just one region, that of the Lake Baykal basin (Figure 1). The region is one with quite high levels of economic activity, yet has a special feature, the unique cultural and historical value of the lake in addition to its high economic value. The need to foster the future economic development of this region and to utilize its riches, on the one hand, and the unquestioned need to preserve the lake, together with its fauna and flora in their original form on the other, therefore does much to explain the important role this region has as a unique model of possible judicious use of natural resources.

There are definite methodological difficulties in depicting the region as a relatively closed economic-ecological system. Even the statement of the problem of depicting the region is contradictory because, as has been pointed out in the foregoing, an intrinsic property of this type of system is openness.

On the other hand, as we understand it, the regional determinant is a known homogeneity of certain criteria within the limits of the region, such as some combination of physical-geographical, economic, administrative, social, and like properties, so that the concept of "boundary of a region" takes on a definite meaning. In our case the problem lends itself to a relatively simple solution. The Lake Baykal basin is a well-defined territory with accurately established boundary, hydrologically speaking. The limits of this territory contain clearly distinguishable landscapes, composed in turn of clearly distinguishable systems of biogeocenoses. And although the landscape and the biogeocenosis are essentially open systems from the energy flow standpoint (sunlight, heat), as well as from the standpoint of their interrelationships with adjacent regions (for, in the final analysis, this is what the evolution of the biosphere is all about), it is important to remember the conception of the unity of biogeocenoses as elementary "atoms" of the biosphere developed by V. I. Vernadskiy and V. N. Sukachev. The fundamental biological flows of matter and energy are locked within the limits of the biogeocenosis, so the biogeocenosis boundary determinant is the condition that the magnitudes of the flows across the boundary be substantially smaller than the internal flows. Thus it is that the colossal Lake Baykal region "ecosystem", which is a system of relatively closed biogeocenoses, is itself relatively closed.

11. The economic subsystem of the region under review is extraordinarily "open." It cannot be considered closed, even as an approximation, because it

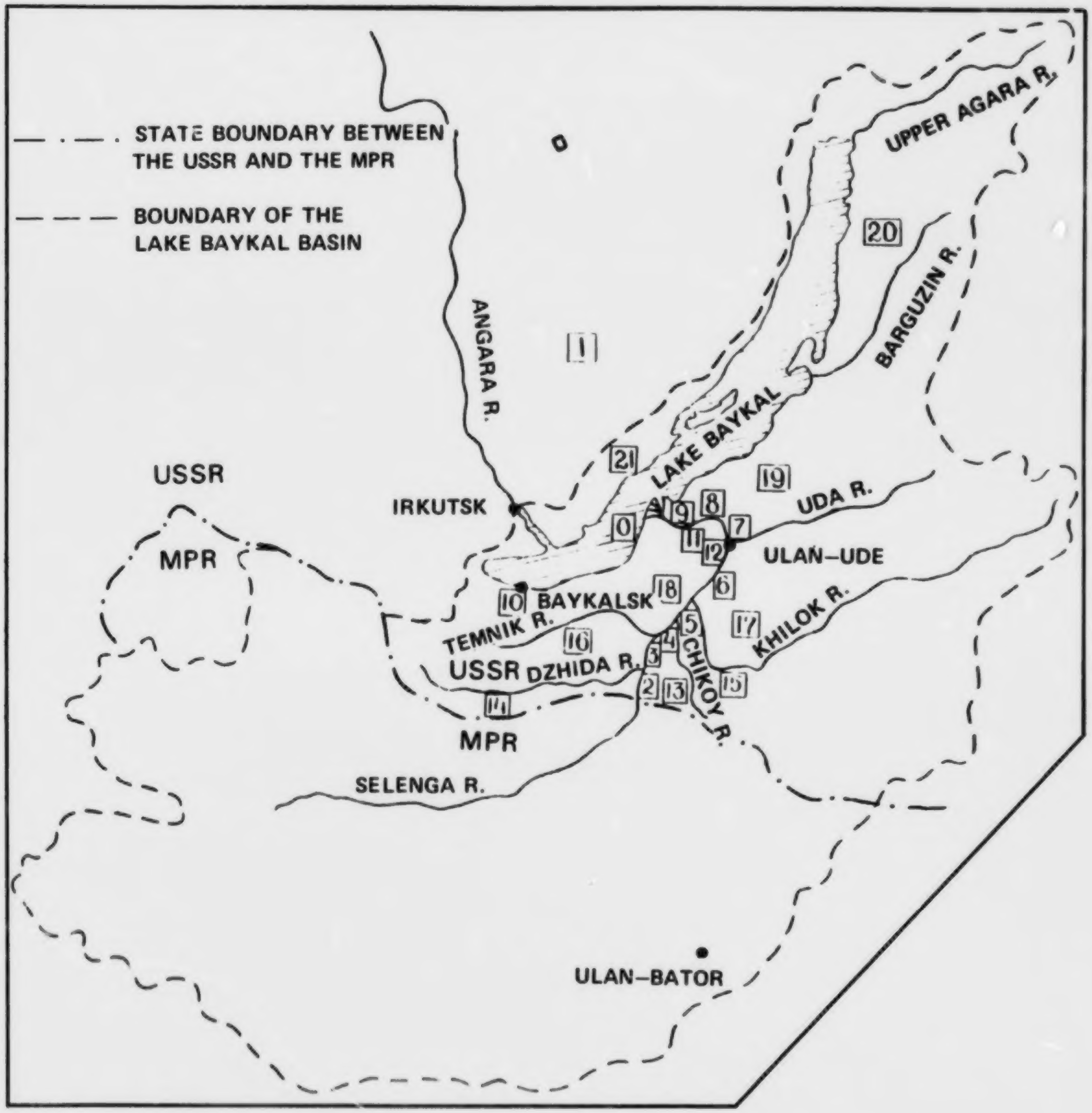

STATE BOUNDARY BETWEEN THE USSR AND THE MPR
BOUNDARY OF THE LAKE BAYKAL BASIN
UPPER AGARA R.
20
BARGUZIN R.
ANGARA R.
1
LAKE BAYKAL
21
USSR
IRKUTSK
19
UDA R.
8
9
0
7
11
12
ULAN-UDE
MPR
10
BAYKALSK
18
6
TEMNIK R.
16
5
17
KHILOK R.
4
CHIKOY R.
USSR
DZHIDA R.
3
2
13
15
14
MPR
SELENGA R.
ULAN-BATOR

FIGURE 1. THE LAKE BAYKAL BASIN AND SELECTED TERRITORIAL SUBSYSTEMS

is obvious the level of economic activity, as well as the region's demographic features (population and its structure), depend most heavily on exogenous factors, while the flows of materials and goods, let us say, across the regional boundary are of the same order of magnitude as are the internal flows.

However, disregarding analysis of the relationships between the region selected and other regions, the exogenous factors mentioned can be presented in the form of deterministic "forces" controlling the development of the economics, and thereby formally close the system.

12. So the region is a relatively closed system consisting of interacting economic and ecological subsystems. Figure 2 lists the main groups of factors describing the behavior of these subsystems (paras. 14–17 will contain brief comments on these subsystems).

We currently are in the process of developing a simulated mathematical model of the functioning of the Lake Baykal region as an economic-ecological

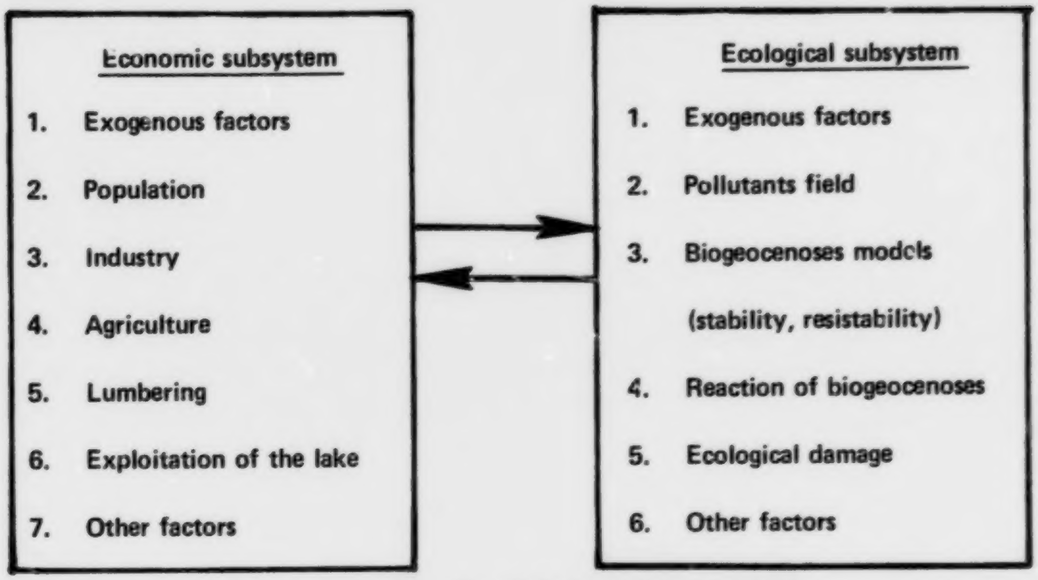


FIGURE 2. TWO OF THE REGION'S SUBSYSTEMS.

system so we can predict the effect economic activity will have on the region's ecological system. We firmly believe that this model, initially of necessity very rough, will be a tool, as well as a powerful stimulus, for future research, because it will help show up the "white spots" (gaps) in our understanding of the functioning of this very complicated system, and thus serve to formulate requirements for the type, quantity, and quality of the needed data, the obtaining of which must be organized. Participation in this work of specialists in the most varied of disciplines who seldom meet as a rule is, in our opinion, extremely valuable. This would include chemists, toxicologists, hydrometeorologists, medical personnel, hygienists, biogeocenologists, industrial workers, and economic geographers.

The development of the model, and full-scale research, forming two continuous, interrelated flows, will become the scientific base for valid decisions on how to standardize anthropogenic pollution on a regional scale.

We cannot provide any details on the work we are doing since it has only just begun, but would, nevertheless, like to share with you some fundamental considerations, and to discuss some of the concrete approaches we have planned for solving the problem.

13. A source of pollution in the regional approach to pollution standardization can be an individual chimney, a plant, a city, or even a whole economic region, depending on how detailed one wants to get, that is to say, what "degree of resolution" of the problem is desired. The sources of pollution in the production-branch approach are individual types of production, or even a branch as a whole. The immediate task confronting us is investigatory in nature, and involves the use of the simulated mathematical model of how the region functions to make a comparison of the different anthropogenic sources of pollution in terms of the degree to which they are harmful to the environment. Our sources of "pollution" will be:

(1) large cities;

(2) industrial plants;

(3) individual, nonlocalized types of economic activity (animal husbandry, fishing, shipping, lumbering, and the like).

A model such as this must be dynamic, naturally, so the comparison will result in approaching a solution to a task of higher order, that of comparing various alternative programs for the economic development of the territorial economic system under review, and making recommendations for the selection of what would, in some way, be an optimum variant. It seems clear that the solution of the problem will involve the need to combine the regional and production-branch approaches if the solution is to be responsible and concrete. The general arrangement of the solution of this problem can be represented in the form shown below (Figure 3).

14. Now let us make a few brief comments on the economic and ecological subsystems in the Lake Baykal region.

The Lake Baykal region's economic subsystem is growing rapidly and is extraordinarily dynamic. We have selected two sectors in the subsystem, the demography and the national economic, for discussion. The group of exogenic factors describes all those external planning and economic "forces," inclusion of which makes it possible to close the system (para. 11).

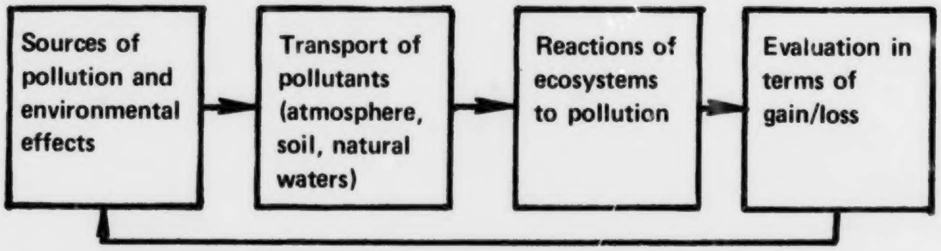

Sources of pollution and environmental effects
Transport of pollutants (atmosphere, soil, natural waters)
Reactions of ecosystems to pollution
Evaluation in terms of gain/loss

FIGURE 3. CLOSED CHAIN PROBLEM SOLUTION

The demographic sector of the economic subsystem can be described in broad outline as follows. Regional development and population are very irregular, and overall are well down on the scale. Mean population density is well below that for the Soviet Union. Immigration from other regions in the country, as well as natural growth, are increasing regional population. The distribution and dynamics of the urban and rural populations are not the same throughout the region, and in fact are determined by the special features of economic development.

The national economic sector describes the region's own economy, the basis of which is well-developed industry (with its timber felling and sawmill branches, as well as branches of heavy and light industry, and power engineering, having an important place in the structure), and diversified agriculture (different branches of livestock breeding, animal husbandry, and plant production).

Hunting and fishing have an important place in the region's economy.

The group of factors combined under the rubric "other factors" includes tourism, and other forms of recreation.

15. So, to deal with the problem we are interested in, that of comparing the different anthropogenic sources of pollution in terms of the damage they do to the environment (para. 13), let us look at the economic subsystem as a whole as a source of pollutants and at its effects on the ecological subsystem, and, in turn, look at the ecological subsystem in terms of its effects on the economic subsystem (Figure 2).

The ecological subsystem is a hierarchy of relatively closed land and water ecosystems (para. 10).

As we know, Lake Baykal is one of the largest fresh-water lakes in the world. Its exceptional age, vast depths, uniqueness of its fauna and flora, and the exceptional purity of its water, all make Lake Baykal stand out sharply from other lakes, and give it its special, unique, character [19].

This very uniqueness, as well as the fact that a huge number (some 400) of rivers empty into the lake (and bring with their waters the pollutants they contain), makes Lake Baykal's biome one of the most critical links within the region. (We should point out that a great many of the natural and cultural values in any region obviously are associated in some way with the natural water systems [3], so that in this respect the Baykal region is a typical example.)

The Selenga is the largest river in the basin. Its drainage basin covers 60% of the total area occupied by the lake's basin. The river is 1024 km in length, with 409 km within the boundaries of the Soviet Union. The upper reaches of the river are in the Mongolian People's Republic.

Only one river, the Angara, flows out of the lake.

16. Irrational management of fishing and sealing (particularly the taking of too many seals, and overfishing the lake without taking reproduction into consideration), and insufficient attention to fish propagation, to say nothing of polluting the rivers with the wastes from the timber industry and with industrial runoffs that cause breeding grounds to deteriorate, can do definite damage to Lake Baykal's biome, and in particular to the fish stocks and to the Baykal seal (ringed seal) population [20].

Change in the lake's chemistry, the result of change in the hydrological and chemical conditions of the lake's most important tributaries, the Selenga and Barguzin rivers, "caused by the development of agriculture and livestock breeding, and by intensification of the feeling of trees, and more recently by the dumping into the river of industrial and domestic sewage," also can harm Lake Baykal's biome [21].

"Change in the composition of organic matter of allochthonous origin attributable to human economic activity such as to increase compounds resistant to oxidation, as well as various organic toxicants that can be the reason for the shift in the equilibrium state of the balance of organic matter in the lake so that bacteria accumulate because their decay is suppressed," is cause for alarm [22].

Efforts have been made in recent years to restock Lake Baykal with fish. A number of farms have been built for the artificial breeding of Baykal omul (*C. autumnalis migratorius*), and steps have been taken to improve natural spawning grounds (timber rafting in free form was prohibited along the Selenga and its tributaries in 1972, for example). Studies of the links between the different ecological and morphological omul groups and the heat and water conditions in the lake, and the unacceptability of establishing strict limits of any sort on fishing for a long period of time, thus become highly important to near-term planning for the reorganization of omul fishing in Lake Baykal, to transferring it along the path of spawning migrations [23, 24].

So far as the ringed seal is concerned (the population of which has been restored and now numbers some 68,000 [25]), it is worth noting that the world has very few lakes with ecosystems similar to the marine ecosystem, those in which highly-organized mammals complete the food pyramid. "The biocenotic and economic importance of the ring seal is that it transforms the huge quantities of small, pelagic, golomanka [*Comephorus*—Baykal fishes related to sculpins] and gobies not now available to man (because they are scattered throughout the body of water and do not concentrate in groups large enough to make their fishing economically advantageous), thereby converting them into a product available in the form of edible and commercial fats, meat, and, in particular, fur skins" [25].

17. Another example of an important ecological link in the region under consideration is the unfavorable change that has taken place in landscapes as a result of economic activity. The problem we inherited from the past, that of moving (eolian) sands and erosion (deflation) of soils in some regions of the Transbaykae is particularly acute, leading as it does to the blocking of populated areas, rivers, roads, to the filling up of lakes and irrigation canals, and covering fields, meadows, and forests. The problem arose many decades ago, and is the result of the destruction of vegetation, brought about by continuous concentrated logging, frequent fires, overgrazing of cattle, as well as ploughing up the light soils of the broad steppes (and thus exposing earlier anchored sands).

Movement of the sands, and the deflation processes, have intensified sharply in recent decades, and this has increased the area of moving sands in the south of the Buryat ASSR several times over. "This outbreak of deflation is merely the reflection of the increasing changes that are taking place in the landscape in the area because of anthropogenic factors" [26].

18. The examples of critical ecological links cited confirm the obvious fact (para. 1) that dealing with industrial and domestic types of pollution is closely interwoven with the broader questions of the rational use of natural resources (land and forest, in particular). The task of comparing different anthropogenic sources of pollution in terms of their "harmfulness," therefore, can be performed only within the framework of a model broad enough to include the questions mentioned. Moreover, the model should contain the whole approach to ecosystems, something that is possible only when problems of a qualitative understanding and of quantitative modeling of the behavior of ecosystems are solved in terms of their common characteristics (stability, capacity to resist near the stability limits, and the like). Then too, there may be individual elements of the ecosystems, such as the omul population, the ringed seal, and others, that could be of interest to us.

The development of this model is a complicated business, and that is why the correct structural organization of the model is so important. This is particularly true of the modular principle of its construction, when the model consists of sectors and subsectors that are relatively independent, and only loosely bound to each other, so that individual parts of the model can be modified without the need for substantial alterations to the rest of the parts.

19. The remainder of our discussion will be purely illustrative in nature because its purpose is to explain the approach, without in any way pretending that the description of that approach is complete, or accurate.

Anthropogenic effects on the Lake Baykal ecosystem reduce to the following types (or combinations of them):

(a) the addition of organic and mineral substances peculiar to the lake (particularly in quantities exceeding their natural fluctuations);

(b) the addition of toxic substances (not peculiar to the lake);

(c) the removal of organic substances (fishing, for example);

(d) change in the makeup of the ecosystem.

Types (a) and (b) can be combined under the rubric of "pollution," and in the main are the side effects of economic activity. Type (c) is one of the forms of economic activity. Type (d) often is the consequence of types (a), (b), and (c), but can be independently important (when new forms are introduced into the ecosystem, for example). Figure 4 shows the paths over which pollution can be carried into Lake Baykal.

20. Elaboration of the scheme shown in Figure 4 requires that individual territorial subsystems within the limits of the region under consideration be isolated. In our case, the criterion for isolating subsystems is known homogeneity of the subsystems by degree of pollution and effect on Lake Baykal's ecosystem.

We have isolated 22 territorial subsystems, including Lake Baykal, within the limits of the region on the basis of the available description of the region as an economic-ecological system (Figure 1 shows the numbers of these territorial subsystems).

Lack of space precludes describing these territorial subsystems in detail. We shall simply point

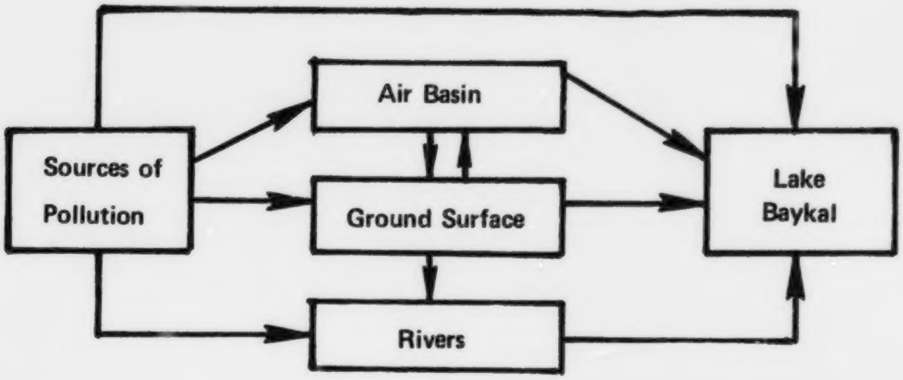


**FIGURE 4. PATHS OVER WHICH POLLUTION CAN BE CARRIED INTO LAKE BAYKAL.**

out the fact that these subsystems include regions that are the largest industrial centers, various hydrological sections (sections of the Selenga River), and various regions around Lake Baykal, and that they are in addition to areas external to the region, and to Lake Baykal itself.

21. Figure 5 shows the interaction between the isolated subsystems as they affect Lake Baykal's ecosystem. Let us explain the subsystems and their interaction (the numbers are those shown in Figure 1):

0. Lake Baykal, the object of the anthropogenic effects in the model under consideration.

1. external areas. The subsystem reflects the overall geophysical transport of pollutants, and transport from adjacent regions.

Fallout directly into the lake, fallout on the watersheds of the rivers in the Lake Baykal basin and subsequent transport into the Lake, and pollutants coming from the upper reaches of the Selenga River (the Mongolian People's Republic), all must be taken into consideration;

2–9. isolated sections of the Selenga River;

10–12. these subsystems describe the effect on Lake Baykal's ecosystems of human settlements and industry in the region. The outlines of regions 10, 11, and 12 (the Baykal Cellulose Plant, the Selenga Cellulose-Cardboard Combine, and the city of Ulan-Ude, respectively) have been drawn to correspond to the transport of pollutants as depicted in Figure 4;

13–21. these are the land areas in the Lake Baykal basin, and were isolated because of a hydrological tie with some particular section of the Selenga River, and other rivers in the basin, as well as because of the relative homogeneity of the type of land use (arable land, haymaking, pasture land) and the extent of forest utilization.

22. Let us, in addition to the territorial breakdown, isolate within the economic-ecological system of the region, certain functional subsystems of interest to us:

(1) human settlements, industry, transportation;
(2) resource utilization (the biome of Lake Baykal, forest and land resources);
(3) recreation;
(4) hydrology, and quality of surface waters;
(5) ecological and economic aspects.

Let us now present the factors that describe qualitatively the interaction of these subsystems in the form of a matrix (Table 1). The matrix element with subscripts i, j (i is the row number, j is the column number) describes the output factors of the $i^{th}$ subsystem with i = j, and these are simultaneously the input factors of the $j^{th}$ subsystem (when i = j, the element presents the concise content of the corresponding subsystem).

As may be seen from Table 1, the presence of diverse (including reverse) interrelationships between these subsystems makes the model extraordinarily dynamic (so the important role played by a methodology that simulates modeling when investigating similar systems becomes quite obvious, as we pointed out in paras. 7, 13).

A model such as this should be "rotated" successively for the individual territorial subsystems (in accordance with Figure 5). The result is an assessment of the effect the various anthropogenic sources of pollution have on Lake Baykal's ecosystem.

23. In conclusion, let us point out that the degree of spatial and/or functional disaggregation we have adopted on a preliminary basis in the model can be insufficient (or excessive, perhaps ?) for reaching substantive conclusions. We are assuming, however, that the modular principle we have adopted for building the model is such that the necessary degree of disaggregation, or aggregation, can be introduced in the future without significant rearrangement of the structure of the model.

**Table 1. Interrelationships Between Functional Subsystems**

| SUBSYSTEMS | 1. Human settlements, industry, transportation. | 2. Resource utilization (biome of Lake Baykal, forest and land resources). | 3. Recreation. | 4. Hydrology, and quality of surface waters. | 5. Ecological and economic aspects. |
|---|---|---|---|---|---|
| **1. Human settlements, industry, transportation.** | Population, level of industrial activity, intensity of transportation. | Logging, plowing, laying roads. | Development of suburban rest areas. | Sources of pollution (domestic, industrial run-offs, pollution by shipping) | Public health, future development of industry, and transportation. |
| **2. Resource utilization (biome of L. Baykal, forest and land resources).** | Possibility of unfavorable consequences (blocking populated areas and roads, shoaling of waterways). | Use of fish resources, sealing, lumbering, land use (hay-making, pasture, arable lands). | Perceived quality of the environment. Sport fishing. | Pollution (eutrophication), soil erosion, clogging of rivers, regulatory effect of forests. | Fishing, shooting of seals, scope of lumbering, pasture and arable land areas. |
| **3. Recreation.** | Development of sanitation institutions, road building. | Man-hours on fishing, hunting; tourism, and the like. | Space-time distribution of recreational activities. | Sources of pollution (coliform bacteria). | Economic advantage, capacity of sanitation institutions. |
| **4. Hydrology and quality of surface waters.** | Possibilities of developing the paper and pulp industry. | Field of concentration of ingredients. | Field of concentration of ingredients. | Transport of water masses and pollutants, self-purification. | Development of the paper and pulp industry, field of concentration of pollutants. |
| **5. Ecological and economic.** | Regulating effect of the ecological and economic aspects on the functional subsystem. | | | | Assessment of the ecological and economic aspects. |

Finally, even the coarse degree of disaggregation we have adopted will, we hope, provide the answer to questions of practical interest.

## BIBLIOGRAPHY

1. Ye. K. Fedorov. *Vzaimodeystviya obshchestva i prirody* [The Interaction Between Society and Nature]. Gidrometizdat., Leningrad, 1972.
2. A. A. Mints. "Geographic Approaches to an Economic Assessment of Natural Resources." V kn: *Ekonomicheskiye problemy optimizatsii prirodopol'zovaniya* [In book: Economic Problems of the Optimization of the Use of Nature]. "Nauka" Publishing House, Moscow, 1973.
3. D. Ehrenfel'd. *Priroda i, lyudi* [Nature and People]. "Mir" Publishing House, Moscow, 1973. David W. Ehrenfeld, Biological Conservation, New York, 1970.
4. R. Waterman. *Teoriya sistem i biologiya* [Systems Theory and Biology]. "Mir" Publishing House, Moscow, 1971.
5. O. Bertalanfi. *Issledovaniya po obshchey teorii sistem* [Research in General Systems Theory]. "Mir" Publishing House, Moscow, 1969.
6. F. Mesarovich. *Issledovaniya po obshchey teorii sistem* [Research in General Systems Theory]. "Mir" Publishing House, Moscow, 1969.
7. B. S. Fleyshman. *Elementy teorii potentsial'noy effektivnosti slozhnykh sistem* [Elements of the Theory of the Potential Efficiency of Complex Systems]. "Sovetskoye Radio" Publishing House, Moscow, 1971.
8. V. V. Nalimov. *Logicheskiye osnovaniya prikladnoy matematiki* [Logical Bases of Applied Mathematics]. Moscow State University Publishing House, Moscow, 1971.
9. I. M. Gel'fand. "Interaction in Biological Systems." *Priroda*, No. 6, 1969.
10. Hamilton et al. *Systems Simulation for Regional Analysis*. MIT Press, 1969.
11. J. Forrester. *Industrial Dynamics*. Cambridge (Mass.), MIT, New York–London, 1961.
12. J. Forrester. *Urban Dynamics*. Cambridge (Mass.), MIT, 1969.
13. J. Forrester. *World Dynamics*. Cambridge (Mass.), MIT, 1972.
14. A. A. Lyapunov. *Problemy kibernetiki* [Problems of Cybernetics], 25. "Nauka" Publishing House, Moscow, 1972.
15. A. A. Mints. *Mathematicheskiye metody v geografii* [Mathematical Methods in Geography]. Kazan', 1971.
16. M. P. Fedorenko. "Economic Problems of the Optimization of the Use of Nature." In the book of the same name, "Nauka" Publishing House, Moscow, 1973.
17. K. G. Gofman. "Methodological Bases for an Economic Assessment of Natural Resources." V kn. *Ekonomicheskiye problemy optimizatsii prirodopol'zovaniya* [In book: Economic Problems of the Optimization of the Use of Nature]. "Nauka" Publishing House, Moscow, 1973.
18. "Basic Assumptions in the Procedure for an Economic Assessment of Natural Resources in Mass Planning and Design Calculations" (project). V kn. *Ekonomicheskiye problemy optimizatsii prirodopol'zovaniya* [In book: Economic Problems of the Optimization of the Use of Nature]. "Nauka" Publishing House, Moscow, 1973.
19. K. K. Votintsev. *Gidrokhimiya ozera Baykal* [The Hydrochemistry of Lake Baykal]. Irkutsk, 1963.
20. B. Buyantuyev. *K narodokhozyaystvennym problemam Baykala* [National Economy Problems of Baykal]. Ulan-Ude, 1960.
21. A. I. Metseryakova. "The Chemical Runoff of Rivers in the Lake Baykal Basin." V kn. *Krugovorot veshchestva i energii v ozerakh i vodokhranilishchakh* [Circulation of Matter and Energy in Lakes and Reservoirs]. Listvenichnoye na Baykale, 1973.
22. K. K. Votintsev, G. I. Popovskaya. "The Circulation of Organic Matter in Lake Baykal." V kn. *Krugovorot veshchestva i energii v ozerakh i vodokhranilishchakh* [In book: Circulation of Matter and Energy in Lakes and Reservoirs]. Listvenichnoye na Baykale, 1973.
23. I. P. Shumilov. "A Procedure for Calculating the Number of Omul That Are Spawned." V kn. *Krugovorot veshchestva i energii v ozerakh i vodokhranilishchakh* [In book: Circulation of Matter and Energy in Lakes and Reservoirs]. Listvenichnoye na Baykale, 1973.
24. N. S. Smirnova-Zalumi, V. V. Smirnov. "The Omul Population in the Ecosystem in Lake Baykal." V kn.

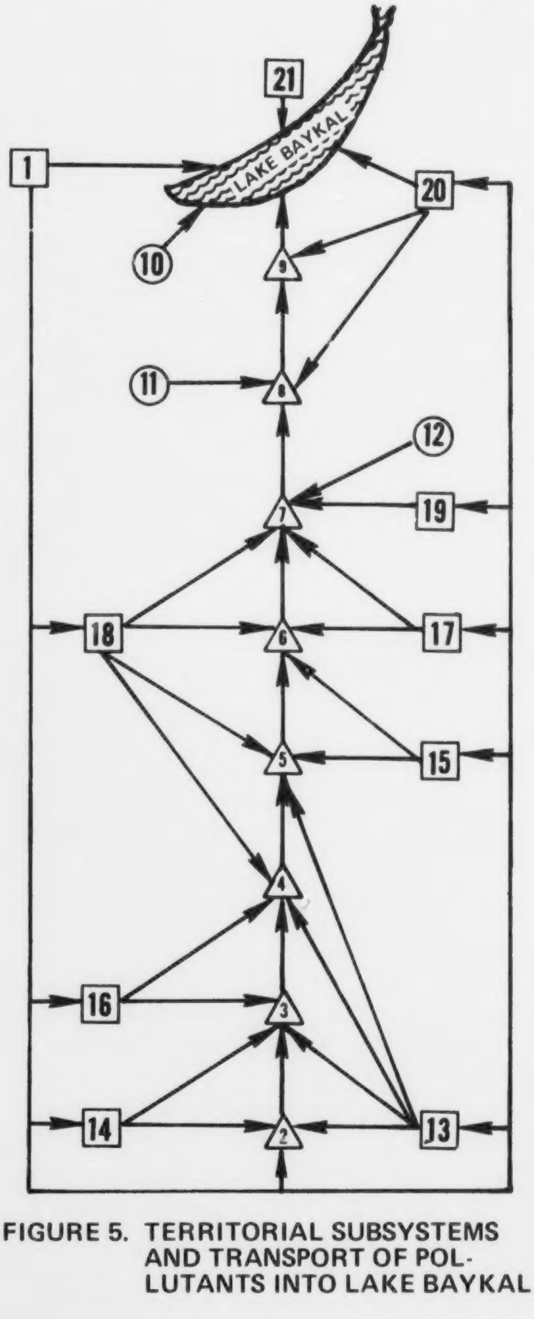


FIGURE 5. TERRITORIAL SUBSYSTEMS AND TRANSPORT OF POLLUTANTS INTO LAKE BAYKAL.

*Krugovorot veshchestva i energii v ozerakh i vodokhranilishchakh* [In book: Circulation of Matter and Energy in Lakes and Reservoirs]. Listvenichnoye na Baykale, 1973.

25. V. D. Pastukhov. "The Role of Seals in the Ecosystems of Continental Water Expanses (the Lake Baykal Example)." V kn. *Krugovorot veshchestva i energii v ozerakh i vodokhranilishchakh* [In book: Circulation of Matter and Energy in Lakes and Reservoirs]. Listvenichnoye na Baykale, 1973.
26. "Wind Erosion of Soils and Measures to Deal With It." *Trudy Buryatskogo instituta yestestvennyikh nauk*, No. 9, 1971. Ulan-Ude.

# A Comprehensive Environmental Analysis of the Upper Potomac Estuary Eutrophication Control Requirements

N. A. Jaworski *

## I. INTRODUCTION

In the United States of America, lakes, rivers, and estuaries are being subjected to increasing threat of eutrophication. Some examples of ecosystems in the U.S.A., that are currently experiencing euthrophication problems are Lake Erie, Lake Mendota, Lake Washington and the Potomac Estuary.

The eutrophication process is a natural phenomenon. However, man has greatly accelerated the rate of eutrophication mainly by the increasing discharge of nutrients from municipal and industrial effluents and the increasing use of agricultural fertilizers.

To comprehensively analyze the environmental aspects of the eutrophication problems in the U.S.A., the following questions are being addressed:

1. What is the current trophic status of the lakes, rivers, and estuaries? How significant is the eutrophication problem?
2. Which nutrient(s) is causing this rate of eutrophication?
3. Can this accelerated rate of eutrophication be decreased? If so, where should the emphasis of the control program be placed?

In this paper is presented a summary of the comprehensive environmental analysis of the eutrophication problem in the upper Potomac Estuary. This case study illustrates the type of analysis that is required to develop a control program for eutrophication. Applicability of this study to other water bodies is also addressed.

## II. BRIEF DESCRIPTION OF THE STUDY AREA

The Potomac River Basin, with a drainage area of approximately 38,000 square kilometers (km$^2$), is the second largest watershed in the Middle Atlantic States. From its headwaters on the eastern slope of the Appalachian Mountains, the Potomac flows first northeasterly and then generally southeasterly some 644 kilometers (km), flowing past the Nation's Capital. The Potomac is tidal from Washington, D.C., to its confluence with the Chesapeake Bay, a distance of 183 km (Figure 1).

The study area includes the tidal portion, which is about 60 meters (m) wide at its uppermost reach near Washington and broadens to nearly 10 km at its mouth. Except for a 7.5 m shipping channel and a few reaches where depths up to 30 m can be found, the tidal portion is relatively shallow with an average depth of approximately 5.5 m.

Of the 3.3 million people living in the entire basin, approximately 2.8 million reside in the upper portion of the Potomac Estuary within the 7,300 km$^2$ which comprises the Washington Metropolitan Area. The lower area of the tidal portion, which drains 8,300 km$^2$, is sparsely populated.

The upper reach above Indian Head, although tidal, is essentially fresh water. The middle reach is normally the transition zone from fresh to brackish water. The lower reach is mesohaline with chloride concentrations near the Chesapeake Bay ranging from approximately 7,000 to 11,000 mg/1.

The average freshwater flow of the Potomac River near Washington, before diversions for municipal water supply, is 305 cubic meters per second (cms) with a median flow of 185 cms. The flow of the Potomac is virtually unregulated and is thus characterized by extremely high and flashy flows otten approaching 2,500 cms during flood conditions and 30 cms during droughts.

## III. WATER QUALITY PROBLEMS

Early historical observations of the water quality conditions include reports that in the later 1790's President Adams swam in the Potomac Estuary near Washington, D.C. By the 1860's when Abraham Lincoln was president, the canals leading into the Potomac Estuary, as well as the Potomac Estuary itself, often emitted objectionable sewage odors forcing Mr. Lincoln to leave the White House at night. From the year 1870, when the first sewers and culverts were constructed, to the year 1938, when the

---

*Director, Pacific Northwest Environmental Research Laboratory, U.S. Environmental Protection Agency

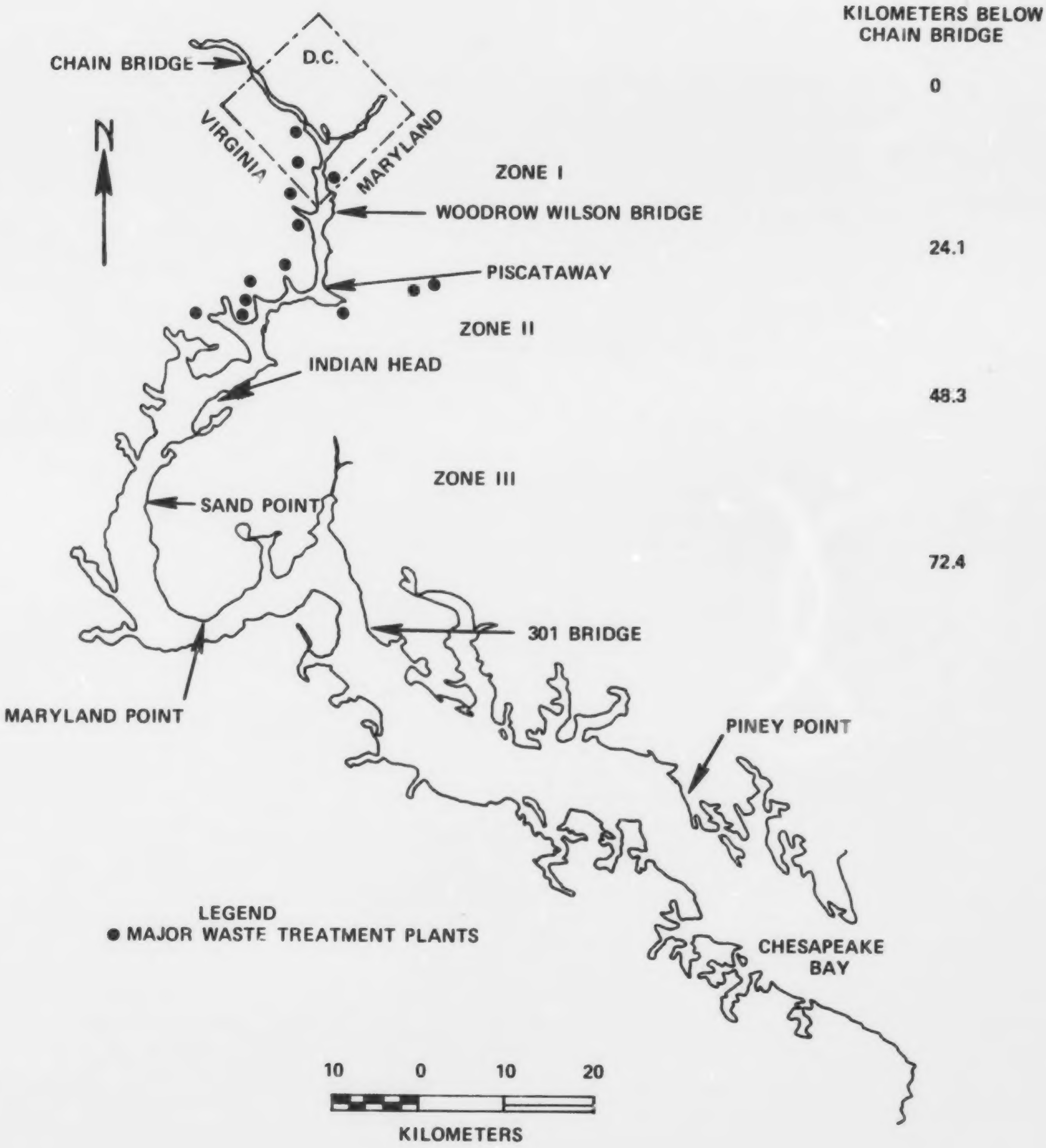

KILOMETERS BELOW CHAIN BRIDGE
0
24.1
48.3
72.4
CHAIN BRIDGE
D.C.
VIRGINIA
MARYLAND
N
ZONE I
WOODROW WILSON BRIDGE
PISCATAWAY
ZONE II
INDIAN HEAD
ZONE III
SAND POINT
301 BRIDGE
MARYLAND POINT
PINEY POINT
CHESAPEAKE BAY
LEGEND
● MAJOR WASTE TREATMENT PLANTS
10
0
10
20
KILOMETERS

FIGURE 1. STUDY AREA

**Table 1**

**Water Quality Problems**
**Upper Potomac Estuary**

| Reach | Kilometers of River Affected | Major Type of Pollution | Major Source of Pollution |
|---|---|---|---|
| Chain Bridge to Hains Point | 11 | Frequently high bacterial counts | Overloaded sanitary sewers and combined sewer overflows |
| Hains Point to Piscataway Creek | 16 | Low-dissolved oxygen concentrations | Effluents from wastewater treatment facilities |
| Piscataway Creek to Maryland Point | 50 | Nuisance algal growths | Nutrients in wastewater discharges |
| Anacostia Tidal River | 13 | Frequently high bacterial counts and low-dissolved oxygen concentrations | Combined and sanitary sewer overflows |

first primary treatment plant was built, almost all of the sewage from the Washington Metropolitan Area was discharged untreated into the Potomac Estuary.

The burgeoning population growth in the Washington Metropolitan Area has compounded the water quality management problem. The accelerated population growth has completely outstripped attempts to provide adequate facilities for wastewater treatment. In addition, much of the growth has been uncontrolled in nature and location, and it is now difficut to provide adequate wastewater collection and treatment within the limited space available for such facilities in the area. Changes in composition of the wastewater, mainly in the phosphorous content, have also had a pronounced effect on water quality.

Since the first sanitary survey was made by the U.S. Public Health Service in 1913 (1), the water quality with respect to bacterial densities and dissolved oxygen levels in the Washington Metropolitan Area has been degraded as a result of the discharge of either untreated or inadequately treated municipal sewage.

The upper estuary has been divided into the four reaches according to type and source of pollution as itemized in Table 1 and shown in Figure 2. There are about 90 kilometers of the upper estuary degraded with the effects of eutrophication being pronounced in approximately 50 kilometers. In addition, the Upper Potomac Estuary, including the Anacostia Tidal River, is subjected to periods of high concentrations of sediment. During initial studies of the estuary, major emphasis was placed on the high bacterial and low-dissolved oxygen problems (2) (3). More recently, the nuisance algal problem has also been included.

The time frame of algal problem development has been developed from several studies as summarized by Jaworski et al. (4). As shown in Figure 3, there have been historical invasions of nuisance growths in the Upper Potomac Estuary.

From a review of data in Figure 3, it would appear that nuisance conditions did not develop linearly with an increase in nutrients. Instead the increase in nutrients appeared to favor the growth and eventually the domination by a given species. As nutrients increased further, the species in turn was rapidly replaced by another dominant form. For example, water chestnut was replaced by water milfoil which in turn was replaced by blue-green algae, mainly *Anacystis*.

The massive blue-green algal blooms, which have occurred every summer since 1960, appear to be associated with large increases in phosphorus and nitrogen loadings in the upper reaches of the Potomac River tidal system (Figure 3). The blooms have persisted since the early 1960's although during this period the amount of organic carbon from wastewater was reduced by almost 50 percent when compared to that discharged prior to 1960.

Under warm temperature and low-flow conditions, large standing crops of this alga develop green mats of cells. Chlorophyll *a* concentrations range from approximately 50 to over 200 u g/l in these areas of dense growth which at times extend over approximately 80 km of the upper and middle reaches of the estuary. These high chorophyll levels are 5 to 10 times those reportedly observed in other eutrophic waters by Brezonik et al. (5) and by Welch (6). During a dense bloom, the dry weight of cells ranges from 10 to 25 mg/l which is almost twice those reported from the lakes in Madison, Wisconsin.

In the mesohaline portion of the lower reach of the Potomac Estuary, the algal populations are not as dense as in the freshwater portion. Nevertheless, at times large populations of marine phytoplankton (primarily the dinoflagellates *Gymnodinium* sp. and *Amphidinium* sp.) occur producing what are known as "red tides."

## IV. NUTRIENT CONCENTRATIONS AND SOURCES

The concentration of nutrients along the estuary varies as a function of wastewater loading, tempera-

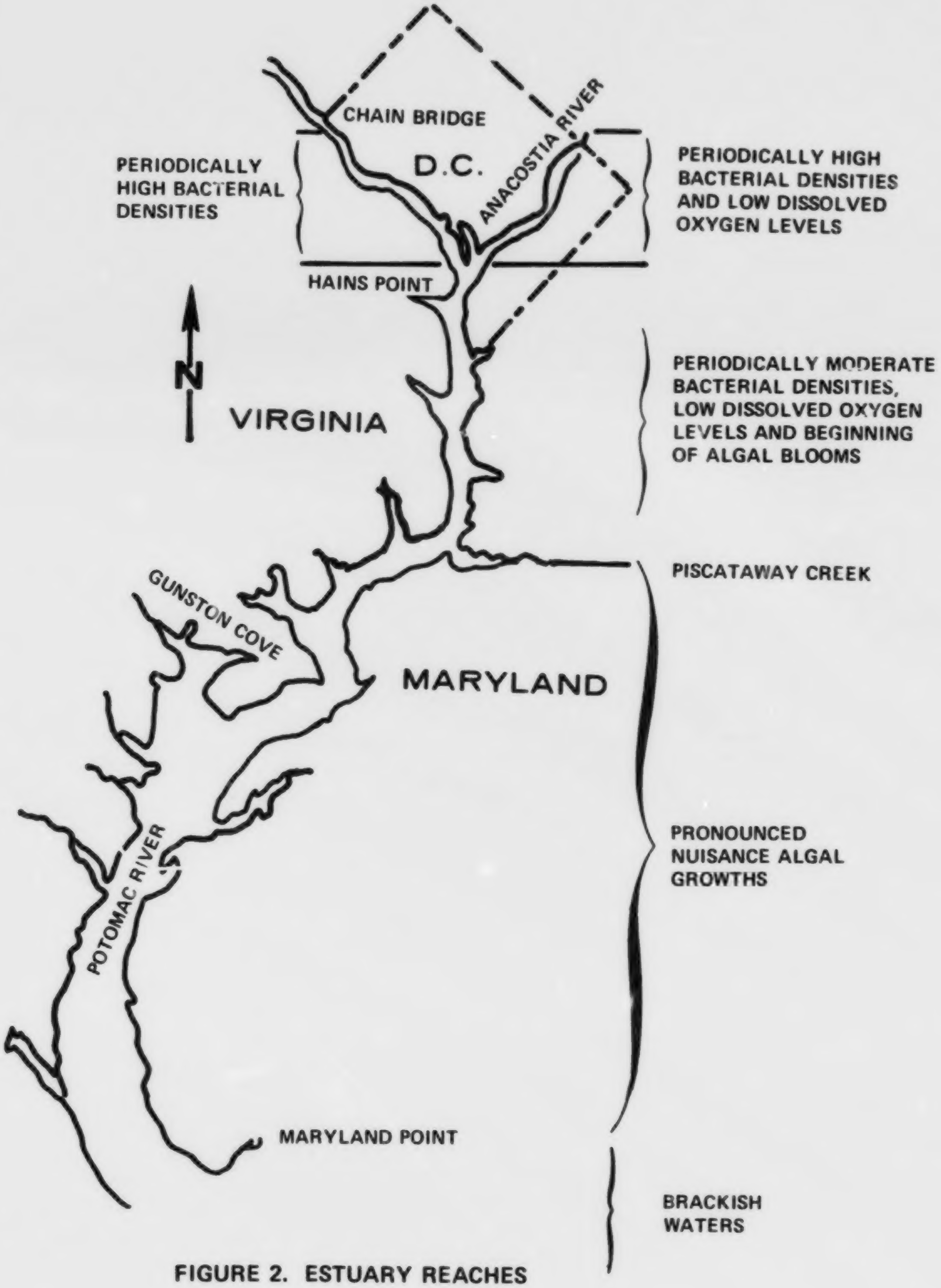

CHAIN BRIDGE
D.C.
ANACOSTIA RIVER
PERIODICALLY HIGH BACTERIAL DENSITIES
PERIODICALLY HIGH BACTERIAL DENSITIES AND LOW DISSOLVED OXYGEN LEVELS
HAINS POINT
N
VIRGINIA
PERIODICALLY MODERATE BACTERIAL DENSITIES, LOW DISSOLVED OXYGEN LEVELS AND BEGINNING OF ALGAL BLOOMS
PISCATAWAY CREEK
GUNSTON COVE
MARYLAND
PRONOUNCED NUISANCE ALGAL GROWTHS
POTOMAC RIVER
MARYLAND POINT
BRACKISH WATERS

FIGURE 2. ESTUARY REACHES

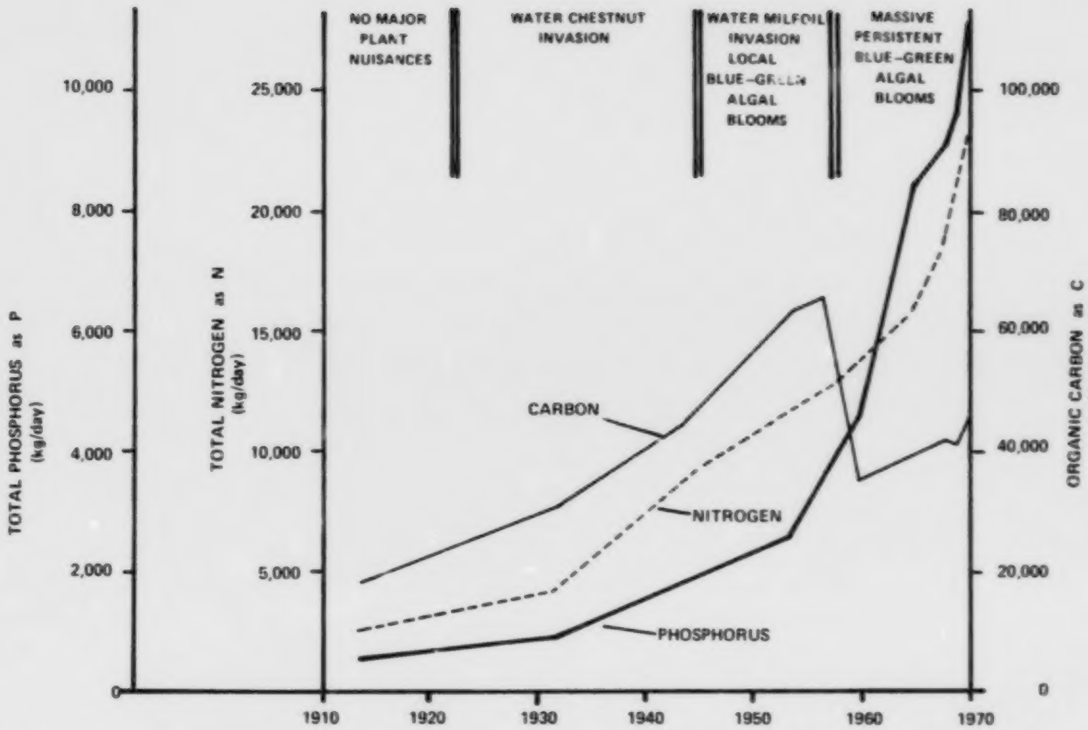


FIGURE 3. ESTUARY HISTORY

ture, freshwater inflow from the upper basin, biological activity, and salinity. The annual distribution of the various nutrient concentrations has been reported by Jaworski, et al. (4), and the summer levels are summarized in Table 2 for five key stations along the estuary.

In the vicinity of the Woodrow Wilson Bridge, there is an increase in alkalinity, total phosphorus, $NO_2$ and $NO_3$ nitrogen, and ammonia nitrogen with a corresponding decrease in pH, all of which can be attributed to the 1.23 million $m^3$ per day of wastewater discharged in the Washington Metropolitan Area. The rapid disappearance of the ammonia nitrogen between Woodrow Wilson Bridge and Indian Head and Maryland Point is caused by the oxidation of $NH_3$ to $NO_2$ + $NO_3$ by the nitrifying bacteria. The sharp drop in $NO_2$ + $NO_3$ nitrogen between Indian Head and Maryland Point is attributable to the large uptake by the pronounced algal growths in this area.

A complete analysis of the nutrient sources in the Upper Potomac Estuary has been made by Jaworski et al. (4). A summary of the major sources is presented in Table 3 for low, median, and high Potomac River flows.

When considering only the upper basin runoff and wastewater discharges to the estuary as summarized in Table 3, it can be concluded that the order of percentage of nutrients controllable by wastewater treatment is (1) phosphorus, (2) nitrogen, and (3) carbon.

While the controllable phosphorus and nitrogen percentages decrease at higher flows, these conditions usually occur during the months of February, March, and April, when temperatures and algal crops are lowest. Since nuisance algal conditions occur primarily in the upper or the freshwater portion of the estuary, the higher flow effects are reduced considerably by the time the blooms are most prolific during the months of July, August, and September.

Under low- and median-flow conditions, both nitrogen and phosphorus are largely controllable. If allowances are made for atmospheric contributions of nitrogen, only an approximate 200 kg/day of nitrogen could be added to the upper estuary, which is less than 10 percent of the nitrogen in the wastewater discharges. Thus, during summer months, algal control by management of nitrogen instead of phosphorus appears to be a feasible alternative.

Using only 0.1 percent of the transfer rate, the amount of carbon ($CO_2$) potentially available from the atmosphere was estimated to be approximately 431,000 kg/day (4). Moreover, with the upper reach of the estuary well mixed due to tidal action,

### Table 2

**Average Range of Concentration**
**Summer Conditions**
**Upper Potomac Estuary**

| Stations and Kilometers from Chain Bridge | pH (units) | Alkalinity (mg/1) | Total Phosphorus (mg/1) | $NO_2+NO_3$ Nitrogen (mg/1) | $NH_3$ Nitrogen (mg/1) |
|---|---|---|---|---|---|
| Chain Bridge (0.0) | 7.5–8.0 | 80–100 | 0.08–0.20 | 0.3–1.0 | 0.10–0.50 |
| W. Wilson Bridge (19.5) | 7.0–7.5 | 90–110 | 0.30–1.20 | 0.8–1.2 | 1.00–3.00 |
| Indian Head (49.3) | 7.2–8.0 | 70–90 | 0.20–0.40 | 0.5–1.5 | 0.10–0.50 |
| Maryland Point (84.3) | 7.5–8.2 | 60–85 | 0.10–0.25 | 0.1–0.3 | 0.05–0.30 |
| 301 Bridge (104.7) | 7.5–8.0 | 65–85 | 0.05–0.20 | 0.1–0.2 | 0.05–0.20 |

### Table 3

**Summary of Major Nutrient Sources**
**Upper and Middle Reaches of the Potomac Estuary**

**Low-flow Conditions**
(95% of time exceeded)
(Potomac River Discharge at Washington, D.C.=40 cubic meters/sec)

| | Upper Basin Runoff * | Percent of Total | Estuarine Wastewater Discharges | Percent of Total | Total (kg/day) |
|---|---|---|---|---|---|
| Carbon | 77,100 | 52 | 72,600 | 48 | 143,700 |
| Nitrogen | 3,000 | 10 | 27,200 | 90 | 30,200 |
| Phosphorus | 450 | 4 | 10,900 | 96 | 11,350 |
| **Median-flow Conditions** (50% of time exceeded) (Potomac River Discharge at Washington, D.C.=185 cubic meters/sec) | | | | | |
| Carbon | 159,000 | 68 | 72,600 | 32 | 231,600 |
| Nitrogen | 18,100 | 40 | 27,200 | 60 | 45,300 |
| Phosphorus | 2,400 | 18 | 10,900 | 82 | 13,300 |
| **High-flow Conditions** (5% of time exceeded) (Potomac River Discharge at Washington, D.C.=1150 cubic meters/sec) | | | | | |
| Carbon | 680,000 | 90 | 92,600 | 10 | 752,600 |
| Nitrogen | 185,000 | 87 | 27,200 | 13 | 212,200 |
| Phosphorus | 10,000 | 47 | 10,900 | 53 | 20,900 |

*Upper basin runoff: both land runoff and wastewater discharges in upper basin.

recruitment of carbon from benthic decomposition appears to be a significant source of inorganic carbon as well. When all potential sources are considered, it appears that management of carbon for algal control is not a feasible alternative at the present time.

## V. EUTROPHICATION CONTROL REQUIREMENTS

For water quality management purposes, the Upper Potomac Estuary may be considered hypereutrophic when nuisance plant organisms become predominant as it is now occurring with the blue-green algae. Four major water use interferences have been offered by Jaworski, et al. (4) including the required reduction in the algal standing crop for each of the conditions as shown in Table 4.

The first two are related to the oxygen budget. Studies have demonstrated that during the summer months more ultimate oxygen demand is added to the upper estuary as a result of these algal growths than from the present wastewater discharges, though this demand may not be fully exerted.

The aesthetic and recreational potential of the upper estuary is impaired by the extensive mats of algae which cause objectionable odors, clog marinas, and cover beaches and shorelines. The potential use of the estuary as a water supply source could also be impaired because of possible toxin problems associated with the blue-green algae.

Of the four interferences, the highest reduction percentages are for control of algal growths to prevent nuisance conditions. From the data in Table 4, a 75 to 90 percent reduction in chlorophyll *a* concentrations will be required to limit chlorophyll levels to approximately 25 $\mu$g/1, the concentration selected as the desired upper limit for eutrophication control in the Upper Potomac Estuary.

**Table 4**
**Subjectve Analysis of Algal Control Requirements**

| Water Quality or Water Use Interference | Indications of Interference | Magnitude of Current Interference * | Desired Limit | Required Percentage Reduction of Current Standing Crop |
|---|---|---|---|---|
| DO Depression Caused by Algal Respiration | mg/l of DO below Saturation | 1.5 to 3.0 mg/l | 0.5 mg/l | 65–85 |
| Increase of Total Oxygen Demanding Load | mg/l of Increase in Ultimate BOD | 15 to 30 mg/l | 5.0 mg/l | 65–80 |
| Recreational & Aesthetic Nuisance Conditions | Chlorophyll *a* Concentration | 100 to > 250 μg/l | 25 μg/l** | 75–90 |
| Toxins | Undefined | Unknown | Unknown | Unknown |

---

*Under nuisance-bloom conditions, chlorophyll *a* concentrations range from 100 to >250μg/l

**Average over entire water column.

## VI. NUTRIENT CRITERIA

The desired nutrient criteria were developed using data from: (1) algal composition analysis, (2) annual nutrient cycles and longitudinal profiles, (3) bioassay studies, (4) review of historical data, (5) comparison with a noneutrophic estuary, and (6) algal modeling. Each method was used independently in the development of a nutrient phytoplankton relationship in the Potomac Estuary.

When investigating the role of nitrogen and phosphorous in eutrophication of the Potomac Estuary, a detailed study of the movement of these nutrients was made using both a real-time dynamic water quality estuary model (8) and an average tidal mathematical model (9). The dynamic model was expanded to predict the concentration of chlorophyll *a* based on the utilization of inorganic nitrogen. In Figure 4, predicted $NO_2 + NO_3$, $NH_3$ and chlorophyll *a* profiles are presented. The predicted maximum concentrations conform closely to observed data in both distribution and magnitude.

From field data, bioassay studies, and mathematical model runs, it was concluded that the standing crop of blue-green algae can be predicted using the nitrogen cycle. This further supports the premise that the nitrogen availability appears to control the standing crop. Similar methods also indicated that if total phosphorus were in the range of 0.03 to 0.1 mg/l, the desired 25μg/l level of chlorophyll could be realized.

Based on the six independent methods of analysis and the 25 μg/l level of chlorophyll *a*, nutrient criteria were developed for reversing the eutrophication process occurring in the freshwater portion of the Potomac Estuary as shown in Table 5. Since there are over 5.0 mg/l of inorganic carbon in the estuary, even under maximum bloom conditions, no criterion for carbon could be established at the present time.

**Table 5**
**Required Concentration for Eutrophication Reversal**

| Parameter | Concentration Range |
|---|---|
| Inorganic Nitrogen | 0.30–0.5 mg/l |
| Total Phosphorus | 0.03–0.1 mg/l |

The lower values of these ranges are to be applied to the freshwater portion of the middle reach and to the embayment portions of the estuary in which the environmental conditions are more favorable toward algal growth. The higher values are more applicable to the upper reach of the Potomac Estuary which has a light-limited euphotic zone of usually less than 0.60 meters.

Since the growth of massive blue-green algal mats is apparently restricted to the freshwater portions and dinoflagellates are often encountered in the mesohaline environment, no specific nutrient criteria have been established for the mesohaline portion of the Potomac Estuary. It appears that if the aforementioned nutrient criteria are achieved in the upper estuary, adequate control of the eutrophication process in the lower reach of the estuary should also be realized.

## VII. WASTEWATER MANAGEMENT ZONES

Initially in 1969, to facilitate the determination of wastewater management requirements, the upper and middle reaches of the estuary were divided into three 15-mile (24 km) zones with similar physical characteristics, beginning at Chain Bridge (see Figure 1). This zoning concept allows for greater flexibility in developing control needs.

More recent studies have suggested that Zone I be divided into three subzones described as follows:

| *Subzone* | *Description* |
|---|---|
| I-a | Potomac Estuary from Chain Bridge to Hains Point, a distance of 12.1 kilometers. |
| I-b | Anacostia tidal river from Bladensburg, Maryland, to the confluence with the Potomac, a distance of 14.4 kilometers. |
| I-c | Potomac Estuary from Hains Point to Broad Creek, a distance of 12 kilometers. |

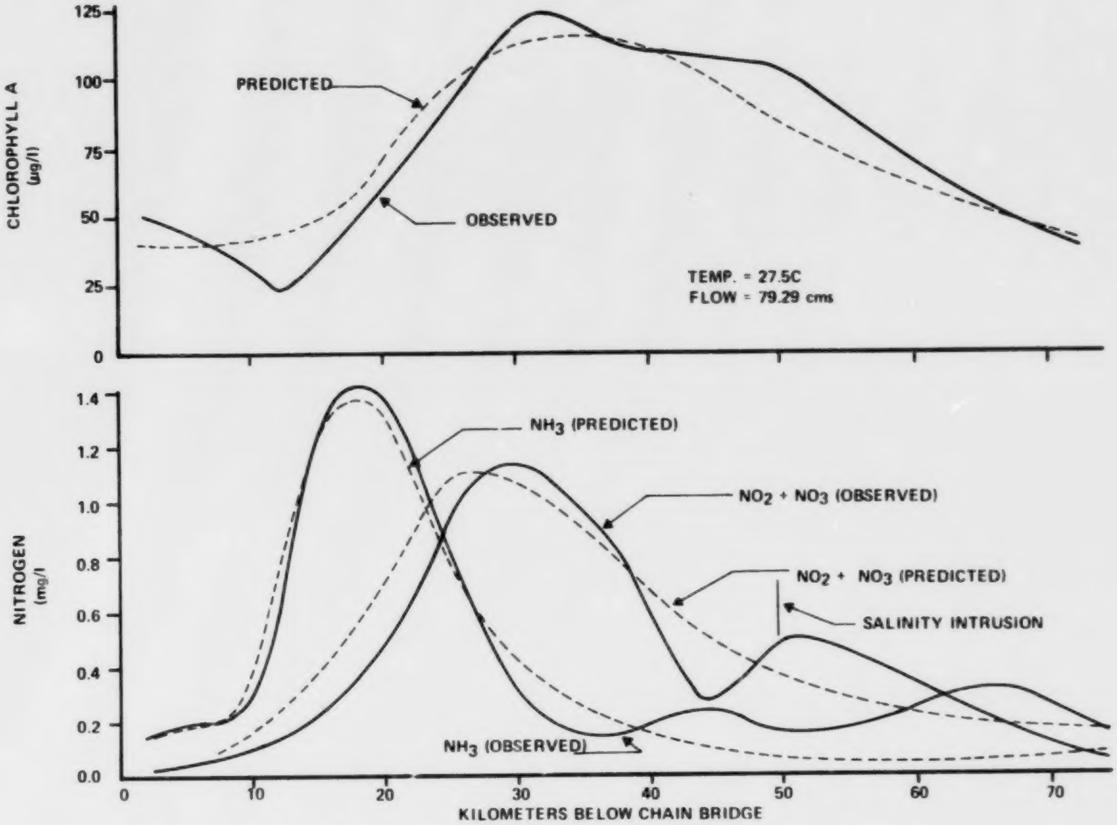


FIGURE 4. PREDICTED AND OBSERVED PROFILES

Discharges into tidal embayments were investigated on an individual basis.

Using the zone concept, total maximum loadings for each pollutant were developed for each zone. Allocation of pound loadings for each discharge can be obtained by prorating the zonal poundage using various bases such as population, drainage areas, geographical subdivisions, and others.

## VIII. WATER QUALITY SIMULATION MODELS

Water quality simulations and wastewater treatment investigations were made using the FWQA Dynamic Estuary Model (DEM) and the DECS III, a general purpose estuarine model. The DEM (8) is a real-time system utilizing a two-dimensional network of interconnecting junctions and channels which permits direct inclusion of tidal embayments in the flow representation. The model is comprised of a hydraulic component that describes tidal movement and a quality component. The DEM includes the basic transport mechanisms of advection and dispersion as well as the pertinent sources and sinks for each constituent. This model was used to simulate water quality conditions on an hourly basis and to determine zonal loadings under low-flow conditions.

DECS III is based on a time-dependent tidal average solution of the basic mass balance equations (9). This model was used to investigate seasonal variations in the nitrogen and phosphorus distributions in the Upper Potomac Estuary.

Simulations of phosphorus discharges into the Potomac Estuary were made using second-order reaction kinetics with a deposition rate of 0.05 mg/day at a temperature of 29°C. The allowable phosphorus loadings were determined based on maintaining an average of 0.1 mg/1 of phosphorus (P) within Zone I, 0.067 mg/1 (P) within Zone II, and 0.03 mg/1 (P) within Zone III.

For investigating the role of nitrogen in water quality management, a feedback system of the introgen cycle was incorporated into the dynamic estuary mathematical model similar to that proposed by Thomann et al. (10). The model consists of six possible reactions: (1) chemical and biological decomposition of organic nitrogen to ammonia, (2) bacterial nitrification of ammonia to nitrite and nitrate, (3) phytoplankton utilization of ammonia, (4) phytoplankton utilization of nitrite and nitrate, (5) deposition of organic nitrogen, and (6) decay of phytoplankton.

## IX. MAXIMUM CONSTITUENT LOADINGS PER ZONE

Using the models and coefficients as described in the previous sections, zonal loadings were determined for UOD nitrogen, and phosphorus (see Table 6). The loadings presented are maximum allowable loadings for each zone, assuming that adjuacent zones are loaded to their maximums.

The increases in loadings for the lower zones mainly reflect the increase in the estuary's volume and tidal transport. Since nitrogen and phosphorus criteria for the lower zones are more stringent, the increase in nutrient loadings in this area is not as pronounced as for UOD.

For the projected 1980 wastewater loading conditions, the anticipated percent removal rates for Zone I-c would be approximately 93 percent UOD, 96 percent prosphorus and 93 percent nitrogen. Since Zones II and III do not currently receive as much wastewater, the removal percentages will not be as high.

## X. SEASONAL WASTE TREATMENT REQUIREMENTS

### *1. Ultimate Oxygen Demand*

The maximum allowable UOD loadings, as presented in Table 5 for the three upper zones of the Potomac Estuary, were developed for low-flow and summer temperature conditions. During high temperature periods, the effects of nitrogenous oxygen demanding substances on the dissolved oxygen budget were determined to be quite significant.

Studies have shown that during very warm periods, when nitrification rates are high, the nitrogenous component of UOD exerts 250,000 lbs/day of oxygen demand as compared to approximately 200,000 lbs/day from the carbonaceous demand. During low temperature periods, when the ambient water temperature is less than 15°C, the effects of nitrification on the dissolved oxygen budget have been shown to be negligible.

Based on these findings, it was recommended that (1) UOD loading presented in Table 5 be applied only under summer conditions, (2) the removal or oxidation of ammonia in wastewater discharges be provided whenever the water temperature is above 15°C, and (3) a high degree of removal of suspended solids (a maximum of 15 mg/1 in the effluent) and carbonaceous oxygen demanding material (a minimum of 90 percent) be provided on a year-round basis to prevent the accumulation of sludge deposits in the vicinity of sewage treatment plant outfalls during cooler weather and to maintain high DO levels under ice cover.

**Table 6**
**Maximum UOD, Phosphorus, and Nitrogen Wastewater Loadings for Low-Flow Summer Conditions**
**(kg/day)**

| Zone | Allowable UOD | Phosphorus | Nitrogen |
|---|---|---|---|
| I-a | 1,800 | 90 | 450 |
| I-b | 1,400 | 40 | 140 |
| I-c | 33,800 | 400 | 1,580 |
| II | 85,500 | 680 | 2,600 |
| III | 171,000 | 900 | 4,100 |

### *2. Phosphorus and Nitrogen*

The loadings, as presented in Table 6, were established for low-flow conditions. During these periods, the nutrient contribution from the upper basin is insignificant when compared to that contained in the wastewater discharges.

To determine whether the nitrogen and phosphorus criteria could be met under varying Potomac River inflows and varying nutrient contributions from the upper basin, an annual simulation was made of conditions from February 1969 to September 1970. This period was critical because a drought condition occurred during June and July of 1969, and August flows were over four times above the average discharge. Thus, both low and high summer flows were simulated.

Mathematical model analysis of the annual distribution of phosphorus in the critical algal growing area showed close agreement between the observed and predicted phosphorus profiles (see Figure 5). Also shown in Figure 5 are the predicted annual phosphorus profiles resulting from year-round wastewater phosphorus removal in the upper estuary, assuming: (1) no control and (2) 50 percent control of the phosphorus loading originating in the Upper Potomac River Basin. From the data presented in Figure 5, it was concluded that both (1) the adherence to maximum allowable phosphorus loadings from wastewater effluents being discharged directly into the estuary (see Table 6) and (2) a 50 percent reduction of the total incoming phosphorus load from the upper basin, will be required if the recommended maximum phosphorus criteria are to be realized. In order to achieve a 50 percent reduction in the present phosphorus load from the Upper Potomac River Basin, the current overall wastewater contribution of 2700 kgs/day must be reduced to less than 320 kgs/day.

Because of the more stringent criteria, particularly in the lower zones including longer transport time, the possibility of recycling previously deposited phosphorus from bottom muds and the unpredictability of phosphorus in various forms being transported from the upper basin, year-round phosphorus removal at all wastewater treatment facilities in the Potomac River Basin was recommended.

As presented earlier, the necessity for unoxidized nitrogen control in wastewater discharges to maintain a high dissolved oxygen content in the Potomac Estuary was restricted to that time of year when water temperatures exceed 15°C. When evaluating the need for annual nitrogen control to prevent excessive algal blooms, controllability, duration of nuisance blooms and temperature become significant factors.

While spring blooms of diatom algal cells have been observed, the major nuisance blue-green algal blooms of algae usually occur during the months of July, August, and September. During these months, the controllability of nitrogen by wastewater treatment is usually the greatest and the water temperature highest.

Mathematical model predictions of inorganic nitrogen concentrations in critical algal growing areas based on (1) no estuary wastewater nitrogen removal, (2) nitrogen removal during periods with temperatures above 15°C (April-November), and (3) year-round nitrogen removal are presented in Figure 6. For the nitrogen loading as given in Table 6, the inorganic nitrogen concentration of less than 0.3 mg/1 can be achieved for drought conditions such as in June and July. The abnormally high August Potomac River flow condition and resulting high upper basin loading caused the nitrogen level to increase to approximately 0.5 mg/1.

While it may be desirable to maintain nitrogen concentrations at or below the selected criteria at all times, the high flows from the upper basin during the winter and spring months contribute high nitrogen loadings which increase the nitrogen concentrations above acceptable levels regardless of wastewater treatment practices. In considering (1) that nuisance algal growths occur mainly during the months of July, August, and September, (2) that seasonal nitrogen removal is generally adequate for maintaining the desired nitrogen concentration during this time, and (3) that unoxidized nitrogen control is required only for warm temperature periods, it was recommended that nitrogen removal for algal control, as in the case of nitrogenous demand for oxygen enhancement, be limited to periods when water temperatures in the estuary exceed 15°C.

In developing the seasonal requirements, emphasis was placed on maintaining a balanced ecological community structure in the upper or freshwater portion of the estuary. More research efforts in both transport mechanisms and nutrient algal relationships are needed to determine management requirements for the lower or saline portion of the estuary.

## XI. ESTIMATED COSTS

The present cost of providing for additional wastewater flows and treatment requirements from the year 1970 to 2020, including operation, maintenance, and amortization cost, has been estimated to be $1.34 billion, with a total average annual cost of $64.9 million. The unit treatment processes assumed include activated sludge, biological nitrification-denitrification, lime clarification, filtration, effluent aeration, and chlorination.

Reduction of the initial capital and operation and maintenance costs to a per capita basis are shown in Table 7. The above summary, which does include replacement cost, indicates that the cost of wastewater treatment in the Upper Potomac Estuary is about $13 to $24 per person/per year. This expenditure, which includes the cost of the activated carbon process, will renovate the water to the chemical and bacteriological levels to meet drinking water quality standards.

## XII. APPLICABILITY OF POTOMAC ECOSYSTEM ANALYSIS TO OTHER EUTROPHIC SYSTEMS

In the summary of comprehensive analysis presented for the Potomac Estuary ecosystem, water quality problems of eutrophication and low dissolved oxygen were addressed. Limits were developed for nutrients and ultimate oxygen demanding materials. For other estuaries along the North Atlantic Coast the pollution stress analysis would be very similar. However, for some lakes and rivers, the comprehensive analyses can be signicantly different.

To cope with this variation aspect of the eutrophication analysis, a National Eutrophication Survey was initiated in 1972. The objectives of the survey are as follows:

1. To assess the current condition of lakes and impoundments (over 800) which have, or are being threatened with, eutrophication problems.
2. To develop scientific rationale which can be used in setting and revising water quality criteria for nutrients and developing eutrophication priorities and policies.
3. To assist federal and state governments in setting nutrient control priorities and policies.
4. To evaluate methods, including analytical techniques, that can be used for lake classification.
5. To assess the contribution of non-point sources of nutrients in varying geographic areas, and, on a limited basis, relate the contribution of nutrients to land use management.

In addition to the above, research efforts in the U.S.A. are being conducted to develop eutrophication controls for freshwater and marine environments. Specific objectives of this research are to:

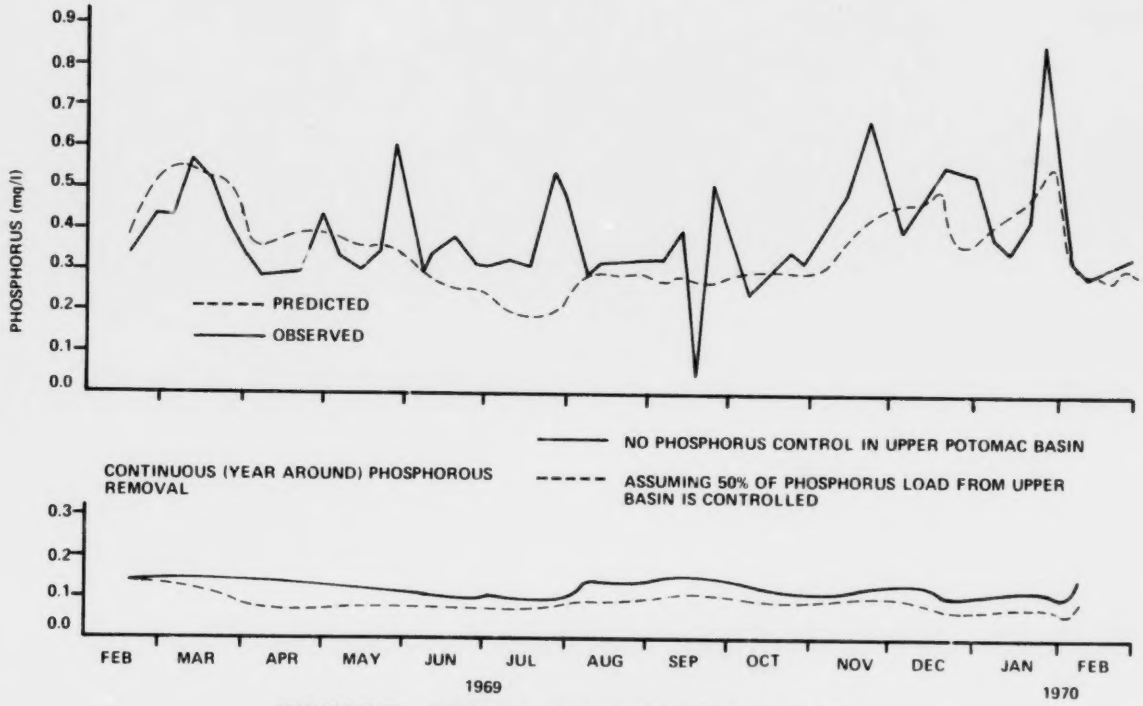


FIGURE 5. ANNUAL PHOSPHORUS PROFILES
POTOMAC ESTUARY AT INDIAN HEAD

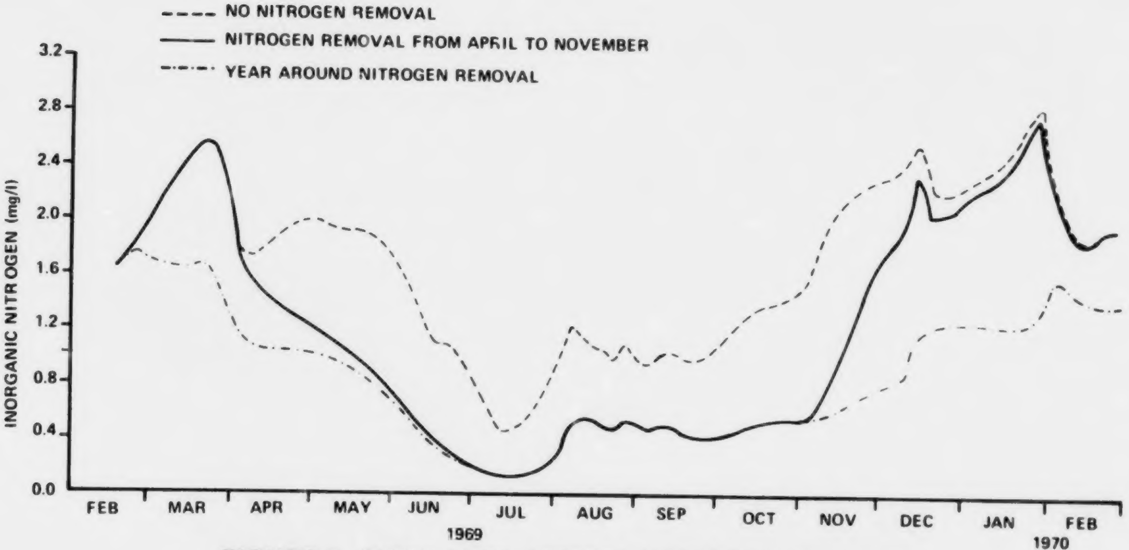


FIGURE 6. SIMULATED ANNUAL NITROGEN PROFILES
POTOMAC ESTUARY AT INDIAN HEAD

**Table 7**

**Per Capita Cost Projections**

| Item | 1970-1980 | 1980-2000 | 2000-2020 |
|---|---|---|---|
| Average Population | 3,350,000 | 5,350,000 | 8,000,000 |
| Initial Capital Cost/Time Period | $570,000,000 | $528,000,000 | $1,173,000,000 |
| Capital Cost/Person/Year | $17.0 | $4.0 | $7.3 |
| O&M Cost/Year | $25,100.000 | $46,200,000 | $72,400,000 |
| O&M Cost/Person/Year | $7.5 | $8.6 | $9.1 |
| Total Cost/Person/Year | $24.5 | $13.5 | $16.4 |

(1) develop an understanding of the eutrophication process, (2) develop methods for monitoring eutrophication conditions and for predicting, with the use of mathematical models, the impact of nutrient sources on eutrophication, (3) develop technology to control and reverse eutrophication processes and (4) to establish the practicability of using this technology through pilot and demonstration scale application.

Our research efforts to date appear to indicate that the approach to comprehensive environmental problems of the Potomac Estuary is adaptable to other ecosystems. However, the uniqueness of many of the lakes and rivers based on morphological and limnological characterstics requires specific analysis and data interpretation. It also appears that control measures will be developed along the morphological and limnological basis for a given set of ecosystems.

## REFERENCES

1. U.S. Public Health Service, "Investigation of the Pollution and Sanitary Conditions of the Potomac Watershed," Hygienic Laboratory Bulletin No. 104, Treasury Department, February, 1915.
2. U.S. Army Corps of Engineers, "Potomac River Basin Report," Vol. I–Vol. VIII, North Atlantic Division, Baltimore District, February 1963.
3. Davis, Robert K., "The Range of Choice in Water Management, A Study of Dissolved Oxygen in the Potomac Estuary," Johns Hopkins Press, Baltimore, Maryland, 1968.
4. Jaworski, N. A., Donald W. Lear, Jr., Orterio Villa, Jr., "Nutrient Management in the Potomac Estuary," Presented at the American Society of Limnology Symposium on Nutrients and Eutrophication, Michigan State University, East Lansing, Michigan, February 1971.
5. Brezonik, W. H., W. H. Morgan, E. E. Shannon, and H. D. Putnam, "Eutrophication Factors in North Central Florida Lakes," Florida Engineering and Industrial Experiment Station, Bulletin Series No. 134, Gainesville, Florida, August 1969.
6. Welch, E. B., "Phytoplankton and Related Water Quality Conditions in an Enriched Estuary," Journal Water Pollution Control Federation, Vol. 40, pp. 1711–1727, October, 1968.
7. Lawton, G. W., "The Madison Lakes Before and After Diversion," Trans. 1960 Seminar on Algae and Metropolitan Wastes, pp. 108—117, Robert A. Taft Sanitary Engineering Center, Technical Report W61-3, 1961.
8. Feigner, Kenneth and Howard S. Harris, "Documentation Report, FWQA Dynamic Estuary Model," FWQA, U.S. Department of the Interior, July 1970.
9. Thomann, Robert V., "Mathematical Model for Dissolved Oxygen," *Journal of the Sanitary Engineering Division, ASCE,* Vol. 89, No. SA5, October 1963.
10. Thomann, R. V., Donald J. O'Connor, and Dominic M. DiTorro, "Modeling of the Nitrogen and Algal Cycles in Estuaries." presented at the Fifth International Water Pollution Research Conference, San Francisco, California, July 1970.
11. Environmental Protection Agency, "National Capital Region Water and Waste Management Report," Washington, D.C., April 1971.

# The Ecological System Stability Concept

V. D. Fedorov *

Complexity must be considered to be the main feature of biological systems. This is so because of the variety of elements comprising the system and the special features of their interrelationships (stress and coordination). The impossibility of studying all the variables of very complex systems inevitably introduces a *subjective* aspect into the selection of some of them, those the specialists deem most important to an understanding of the nature of the phenomenon under investigation. The a priori ranking of system variables according to importance (made consciously or unconsciously) therefore is necessary before the investigation of biological systems of any complexity can be undertaken.

The imposed necessity of operating with a limited number of variables when studying the integral properties of very complex systems precludes the establishment of simple cause-effect types of relationships between the different properties of biological systems. This is what has given rise in biology to the point of view that the *whole* is greater than its component parts, so that the properties of the whole cannot be studied effectively after the whole has been broken down into parts to study the properties of the latter. This approach has led to the concept of the existence of special biological laws that cannot be explained by resort to the positions of the strict propositions of such sciences as physics and chemistry. This point of view is quite productive today, because research in the integral properties of biological systems that reflects the face of the systems, their appearance, fate, and coordination, has been moved up on the agenda.

## 1. *Characteristics and types of biological systems*

Let us define a biological system as an object capable of forming negative entropy, the integral properties of which are the result of the interaction of the parts of which it is formed. Let us now distinguish the following integral properties as primary in biological systems:

(a) complexity, determined by the diversity of the elements involved, and their interrelationships, as well as the stress on the interrelationships in the system;

(b) autonomy, which determines the special features of the position of the system at subsystem levels within the limits of a larger system;

(c) reliability, which determines the probability of processes taking place in the system, and of the temporal maintenance of the constancy of its properties;

(d) stability, which determines the capacity of biosystems to withstand external disturbing effects, while at the same time preserving the system's "face," which is determined by its past history.

There are objective reasons for the subdivision of biological systems into three groups according to the differences in the system of regulations governing their behavior (Fedorov, 1972):

(1) a-type biological systems with central control; that is a subsystem whose regulatory link with any element is stronger than the link between such element and any other, taken arbitrarily. This category includes all animal organisms with central and vegetative nervous systems;

(2) p-type biological systems with "passive" control (Milsum, 1964), which can include plant organisms with their humoral regulation;

(3) r-type biological systems with "random" control, which can include all ecological systems not carriers of a special behavior "program" in the DNA molecules, as is the case in the first two groups.

It is only natural that the importance of the integral properties listed will be dixerent for difierent types of biological systems.

Thus, such integral properties as *reliability* and *autonomy*, on which the carrying out of the "program" is primarily dependent, will be of primary importance for systems with a "program" (a-type and p-type). At the same time, such integral characteristics as *complexity* and *stability*, on which their fate is primarily dependent, that is, the capacity to withstand random "unprogrammed" environmental disturbances, will be of primary importance for ecological systems (r-type).

It is useful, for purposes of convenience in justifying methods used to measure the integral properties of ecological systems, to separate the following from among the "unprogrammed" disturbances of the abiotic component.

---

*Moscow State University, Moscow

Environmental factors that interact with the living component of the system, that is, that have an effect on the system and experience a response from the organisms, should be included among the intrasystem factors (biogenic elements, for example).

The second group of environmental factors is made up of so-called *extra-system* factors. These establish the conditions for the existence of concrete ecosystems within the limits of larger system formations, that experience no noticeable response from the biotic component (solar radiation, temperature, humidity, and salinity, for example).

Finally, the third group of factors is made up of the so-called *disturbing* factors, those for which the ecosystems are poorly prepared (or not prepared at all) by the preceding course of the evolution of the species, and by the past history of the formation of concrete communities under concrete biotope conditions. They may include physically heterogeneous factors, but their distinguishing feature is the suddenness (on the geological time scale) of the effect created by change in their intensity which is unusually abrupt for the system. This action can include an abrupt change in the reserve of any "normal" chemical element in the locus (for example, N and P, which give rise to the problem of eutrophication), or of a normally rare element (for example, salts of heavy metals Hg, Pb, Cd) or, finally, the appearance of a substance that is "new" to life, that has been synthesized by man (for example, DDT, PCB and other pesticides).

In fact, it is indeed the intensification of the disturbing factors associated with the variety of forms, and with the reinforcement of the scales of human activity in the biosphere, that does much to explain just how important the problem of studying the stability of natural systems is today. Therefore, our subsequent elucidation of the material will be limited to a discussion of certain of the general aspects of ecological system stability, and to justifying methods used to measure that stability.

### 2. *Measuring the stability of variables*

A biological system has numerous characteristics that reflect the special features of its organization. These characteristics can be considered as independent variables of the system, and the need arises in any study of a system to continually estimate their behavior in time and/or space.

The measurement of the stability of an independent variable ($S_i$) can be assumed to be the estimate of the variation in the unit values of a function around a mean in order to estimate the temporal change in individual independent variables. Stability thus is a measure of the temporal inconstancy of any function. Such estimates of $S_i$ can be (Fedorov, Bokolova, 1972) the mean values of the deviation in amplitude, $\Delta_1$ [1], or the square of this deviation $\Delta_2$ [2], of the variable with respect to its mean value

$$\Delta_1 = \frac{\frac{1}{n}\sum_{j=1}^{n}\left|\Delta x_j\right|}{\bar{x}}; \qquad (1)$$

$$\Delta_2 = \frac{\frac{1}{n}\left(\sum_{j=1}^{n}\Delta x_j^2\right)^{\frac{1}{2}}}{\bar{x}}, \qquad (2)$$

where

$x_j$ is the value of the variable being measured at the time of measurement;

n is the number of measurements when $\Delta x_j = x_{j+1} - x_j$ and

$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$.

The index of heterogeneity of criterion distribution that we have proposed (Fedorov, 1973), and which is an estimate of the ratio of the mean geometric to the mean arithmetic value of the variable, $\Delta_3$, [3],

$$\Delta_3 = \frac{\left(\prod_{j=1}^{n} x_j\right)^{\frac{1}{n}}}{\bar{x}} \qquad (3)$$

is yet another method for estimating $S_i$. But regardless of the method used to estimate the stability of the individual variables, it can be shown that the set of estimates obtained forms a bimodal distribution (Fedorov, Sokolova, 1972) when the values found for $S_i$ are ranked. One of the sets of estimates here includes the indices that characterize the rate at which the process occurs (that is, the intensity of the function), so these estimates can be expressed as a time derivative. Let us call them functional estimates, assuming that

$$S_f = \frac{1}{n}\sum_{i=1}^{n} S_i \qquad (4)$$

The second set groups the estimates of indices that can be taken to be some result of the processes (functions) at the time of observation, so these estimates can be expressed as a time integral. Let us call them structural estimates, assuming that

$$S_s = \frac{1}{n}\sum_{j=1}^{n} S_j \qquad (5)$$

The most important functional characteristic of an ecosystem is its productivity. The latter reflects the rate of formation of the organic matter (biomass) by the individual species populations making up the community. The most important structural characteristic of ecosystems is the biomass of the populations.

This raises a question. How can the integral properties of a system be evaluated quantitatively from numerous individual estimates of independent variables? Separation of the indices of the ecosystem's biotic component into functional and structural is quite productive in finding an answer to the question just posed because one can *compare* the mean estimates (or modal class positions) for the two sets of estimates obtained, that is, for $S_f$ and $S_s$.

It is readily apparent that comparison of the $S_f$ and $S_s$ estimates leads to the introduction of a certain *measure* that characterizes the *internal* state of the system, and this measure, as will be shown below, determines the homeostasis of the ecosystem.

### *3. The homeostasis of biological systems*

The term *homeostasis* was introduced initially in 1932, by Walter Cannon (Cannon, 1932) to designate the mechanisms that maintain the equilibrium of the internal medium (blood sugar constancy, and constant body temperature in warm-blooded animals, for example), which latter is constantly being upset by changing conditions in the outside world and is constantly being restored by special internal mechanisms with feedback action. These mechanisms are under the control of the autonomic nervous system, so are inherent in all living systems with a central control subsystem; that is, a-type systems, or simply a-systems.

Much later (Fedorov, 1970a), while analyzing the organizational features of a- and r-systems, we turned our attention to the existence of a principle difference between them. Thus, functional reconstructions of a-systems are designed to maintain, in the final analysis, constancy of the structural characteristics (preservation of the specific external appearance of the systems attributable to the functioning of numerous mechanisms for regulating metabolism, for example). Structural reconstructions of r-systems were designed to maintain the constancy of the functional characteristics. Change in the ratio of the photophilic to thermophilic forms, and of the thermophilic to psychophilic species makes it possible to maintain the constancy of the production characteristics during the vegetative period when there is a variation in temperature conditions and illumination. Thus, the presence of ecological mechanisms that periodically maintain the constancy of the functional characteristics of r-systems within predtermined boundaries makes it possible to apply the concept of homeostasis to them after the concept has been generalized and formalized.

We will take it that the concept of homeostasis can be extended to include the mechanisms that maintain the constancy of certain characteristics because of the inconstancy of others. This means that the occurrence of relatively stable properties in one part of the system can be provided because of the occurrence of additional instability in other of its properties.

Thus homeostasis is a regulation mechanism, one that provides for temporal ordering of change in system properties toward stability of groups of characteristics with respect to processes (r-systems), or to their results (a-systems). In ecological systems in particular, the occurrence of an additional zone of inconstancy is associated with change in the relative abundance of individual populations; that is, with change in the specific percentage of biomasses of species while retaining the constancy of the overall formation rate by organisms of a homotypic association.

The relationship

$$G = \frac{S_s}{S_f} \qquad 6$$

was used as the measure of the homeostasis (G) of the living component of an ecosystem in the paper written in collaboration with S. A. Sokolova (Fedorov, Sokolova, 1972). However, for reasons of purely practical convenience associated with introducing the normalization $O \leq G \leq 1$, the measure of the homeostasis is best calculated from Eq. (7) if the value of $S_i$ is found by using Eq. (3) when using the stability index

$$G = 1 - \frac{S_s}{S_f} \text{ when } S_s \leqslant S_f \qquad 7$$

and when calculating the stability estimates using Eqs. (1) and (2)

$$G = 1 - \frac{S_f}{S_s} \text{ when } S_s \geqslant S_f \qquad 8$$

Thus, in a living system a *different* absolute measure of variability can be attributed to a different variable, depending on the method selected for estimating $S_i$.

This gives rise to a routine question. What is the *cause* of the variability in individual characteristics? There are two such causes.

1. First, there are the *internal* causes attributable to the special features of the interaction of the elements comprising the system. These reflect the laws of coordination of elements and of the interrelationships between them. Included, for example, are the variations in the numbers of two populations in a

system whose mutual relations are of the "prey-predator" type. The most important feature of the changes attributable to internal causes is their fluctuating nature; that is, their change with variable amplitude and frequency within the limits of some region that is the *norm* under the system's usual conditions of existence.

2. Second, there are the *external* causes, attributable to the special features of the action of disturbing environmental factors. These are capable of bringing about deviations in system properties specified by the variables that will be beyond the limits of the norm.

Thus, it is precisely the disturbing causes that raise the question of the permissible measure of deviation from the norm, within the limits of which the ecosystem will be able to withstand external actions because of its property of eliminating their consequences as a result of the ordering state of affairs of the homeostasis mechanisms. This capacity is what brings us to the concept of *stability* as an integral characteristic of systems, expressed by the capacity to withstand disturbances. Consideration of stability as the maintenance of some norm for the properties, attributable in turn to the norm for the individual variables, permits posing the question on the plane of establishing a permissible measure of deviations from the norm of certain of its properties under the effects of concrete environmental disturbances.

*Then let us define system stability as the permissible (without risk of destroying the system) measure of deviations in specified properties of a system from the norm, caused in some measure by external disturbing actions.* In this definition, system stability is established first with respect to a limited number of selected properties, properties which we will call the specified properties, and second, with respect to a series of concrete disturbing environmental actions. System stability in this case can be judged *in proportion to* the deviations from the norm of the specified properties caused by the disturbing *measure* of the external action. This approach permits us to use the quantitative estimates of stability to compare estimates characterizing, on the one hand, the measure of the disturbing actions and, on the other hand, the evoked measure of deviations of the specified properties from the norm. This then gives rise to the need to establish not simply a measure of deviation of specified system properties but the *permissible* measure of the deviations which, in time, can be eliminated by the system itself. This measure of deviations eliminated by the system is in fact the permissible measure, and this measure must be considered the measure of stability of the *entire* system as a whole.

In concluding this section it seemed important to us to turn our attention to the idea of the link between the *permissible measure of deviations* in the specified properties of a system and the *measure of the homeostasis*. The very idea of such a link is extremely constructive, because the measure of the homeostasis was quite uniquely established above, whereas the question of the *permissible* measure of deviations (as distinct from the idea of a link with the measure of the homeostasis) immediately acquires acuity and the vagueness of the question concerning "norm" and "pathology."

### *4. Invariants as indices of the stability of biosystems*

A conclusion as to the instability of systems does not follow, of course, from inconstancy of individual variables (estimates of $S_i$). The already generalized interpretation of the homeostasis logically leads to the invariant possibility. It is true, let us stipulate, that an estimate in the form of a mean [Equations (4) and (5)] cannot be invariant, strictly speaking, because here we are estimating some mean instability of individual indices which are incapable of reflecting the more integral characteristics of system *state*. However, as a very rough approximation, when the Condition $O \leq G = \text{constant}$ is satisfied, the simplest primitive form of what we have decided to consider in the form of an invariant turns out to be the same mean estimate of the stability of the functional indices of an ecosystem.

A more typical case of an invariant, when the *properties* under consideration in the norm are generally temporally unchanged (or invariant), can be attributed to a complex combination of consistently varying individual variables. The definition "consistently varying" implies a strict observance of the rules for the coordinating of functions, which has been achieved using the autonomic nervous system in systems with central control, as a result of which it is easier to find invariants for a-systems than for r-systems.

We thereby arrive at a more *integral interpretation of stability (although less constructive) as the temporal maintenance of invariant properties, thanks to the combined, ordered cooperation of continually changing individual system variables.*

Let us define the conditions needed for the manifestation of invariant properties in biosystems. There are two such conditions:

(1) the closed nature of the cycle of the elements of a stagnant environment in ecosystems; (2) the constancy of the instantaneous rate of change in system entropy.

Then, in accordance with these conditions, the invariants of ecological systems can be:

(a) the constancy of the ratios between the synthesis and decay reactions of organic matter in the system (in units of energy);

(b) the constancy of the ratios between the rates of movement of individual elements through the links of the trophic chain, or even of the rate of individual elements;

(c) the constancy of the ratios of negative entropy, forming at different trophic levels in the system.

However, the ecosystem invariants cited and predicted on the basis of general considerations are difficult to use in practice because the determination of their values requires a great deal of preliminary experimental work, and both necessary conditions are almost never satisfied, or are not strictly satisfied.

The concrete generalized indices that can be measured can be (a) *within the limits of a single trophic link,* the BR* index (Fedorov, 1972), where B is the biomass, and R is the rate of its formation, or (b) *within the limits of neighboring trophic links,* the $B_pB_v$ index, which reflects the interaction between predator ($B_p$) and prey $B_v$).

Equations based on BR* and $B_pB_v$ estimates can be derived and the following biologically reasonable assumptions can be useful as the initial premises:

$$R^*_1 \cdot B_1 + R^*_2 \cdot B_2 + \cdots R^*_k \cdot B_k = \sum_{i=1}^{k} R^*_i \cdot B_i = \text{const} \qquad (10)$$

$$\frac{R^*_1 \cdot B_1}{R^*_2 \cdot B_2} = \frac{R^*_2 \cdot B_2}{R^*_3 \cdot B_3} = \ldots = \frac{R^*_k \cdot B_k}{R^*_{k+1} \cdot B_{k+1}} = \text{const} \qquad (11)$$

$$\left[B_p^1 \cdot B_v^1 + B_p^2 \cdot B_v^2 + \ldots B_p^k \cdot B_v^k\right] : K = \text{const} \qquad (12)$$

$$\frac{B_p^1 \cdot B_v^1}{B_p^2 \cdot B_v^2} = \frac{B_p^2 \cdot B_v^2}{B_p^3 \cdot B_v^3} = \cdots = \frac{B_p^{k-1} \cdot B_v^{k-1}}{B_p^k \cdot B_v^k} = \text{const} \qquad (13)$$

What must be emphasized is that the BR* and $B_pB_v$ indices are not system invariants in the full sense of the word. We can only state that they reflect, in implicit form, the movement of energy in the system, and therefore can be used. It therefore is more correct to call them (as well as the biomass and productivity indices which form them, incidentally) system *variants,* postulating that the complication of the variants, justified by logical premises, will bring them asymptotically close to the theoretically possible invariants of an ecosystem. A logically justified example of just such a complication of variants can be the indices BR*D and BR*D/lg W, where D is the information index of diversity (MacArthur, 1955) and W is the number of species.

### *5. Measuring ecosystem stability*

Once the basic concepts included in the definition of ecosystem stability have been defined and concretized, we can move on to discussion of methods of estimating its total stability. Diverse, and numerous, methods for this kind of estimate can be devised, of course. Several estimating methods that use the so-called coefficient of stability, which we shall designate $\chi$, are suggested below.

a. *The canonical method.* If we also attribute the possibility of deviation from the norm to the disturbing source, then the measure of the disturbing action turns out to be precisely the measure of the deviation of an independent variable from the norm, that is,

$$\frac{|x|_n - |x|_1}{x_n} = \frac{\Delta x}{x_n} \qquad (14)$$

Then the result of the action, or the system's response according to the action in terms of the specified properties, $\Delta x$, may be expressed by the relative deviation of a dependent variable from the norm, that is,

$$\frac{|y|_n - |y|_1}{y_n} = \frac{\Delta y}{y_n} \qquad (15)$$

The ratio of these magnitudes can be called the coefficient of system stability

$$\chi = \frac{\Delta x}{x_n} \Big/ \frac{\Delta y}{y_n} = \frac{\Delta x \cdot y_n}{\Delta y \cdot x_n} \qquad (16)$$

In the ecosystem "norm," $\chi$ will equal O when the "norm" of the action of the external factors is preserved, as well as when the system's invariant properties are maintained by the mechanisms of homeostasis.

b. *The selective method* of estimating the coefficient of stability is based on the taking into account of the inhomogeneities in the distribution of the estimates of the actuating source and the response registered by the system [see Eq. (3)].

Here it is assumed that the index of stability $S_{1(\Delta_3)}$ of the variables being measured and the system properties under consideration (variant) is a suitable measure for estimating inhomogeneity in the sets of values for the individual variables being compared. In other words, an implicit identity between the method of estimate $\Delta x$ and the expression $(1 - S_{1(\Delta_3)})$ is assumed. This method avoids the need for tedious work in estimating norms and measures for permissible deviations in an ecosystem. Then

$$\chi = \frac{[1 - S_{(\Delta_3)}]_{res}}{[1 - S_{(\Delta_3)}]_{act}} \tag{17}$$

where $S_{(\Delta_3)}$ is the index of inhomogeneity in the distribution of the estimates of variables, as calculated using Eq. (3). It is important to note that the estimate for one of the variants to which we attribute decisive importance when determining system state (for example, the estimate in terms of BR*D), serves as the system response in this case. At the same time, the estimate of the stability of *only* one of the independent variables, to whose action we attribute system response, also can be used as the disturbing source.

c. *A variant of the selective method* is the estimate of the coefficient of stability in terms of the mean values of the set of estimates, $S_i$, equated to the disturbing sources and to "system responses"

$$\chi = \frac{1 - \frac{1}{n}\sum_{1}^{n} S_{i\,(res)}}{1 - \frac{1}{n}\sum_{1}^{n} S_{j\,(act)}} \tag{18}$$

Let us call this variant the *exploratory method*, one used in studying the combined action of disturbing sources on an ecosystem.

d. *The three means method* uses the mean values of the stabilities, determined by any of the methods relating to the estimates of the disturbing source $(\bar{S}_d)$ and to system state, $\bar{S}_f$ and $\bar{S}_s$, to estimate the coefficient of stability,

$$\chi = \frac{G}{1 - S_{(\Delta_3)}} \tag{19}$$

If the equivalance of $1 - S_{(\Delta_3)}$ and $S_d$ is assumed, a variant of the three means method is the estimate

$$\chi = \frac{G}{S_d} = \frac{1 - \frac{S_s}{S_f}}{S_d} = \frac{S_f - S_s}{S_d \cdot S_f} \tag{20}$$

*6. Conclusion*

In discussing the representation of stability that has been developed, let us turn our attention to what are, in our opinion, several pertinent features that give our constructions a concrete nature.

I. The first of them concerns the division of the characteristics of an ecosystem into functional and structural, so that a quantitative measure of the *homeostasis* (G), can be introduced.

II. The same variable, but used in different recording methods, can simultaneously be the basis for obtaining the characteristics of the disturbing source (for example, an increase in the total reserve of nitrogen in the system, $\Delta N$, leading to eutrophication), a structural characteristic (the content of different forms of nitrogen at the current moment in time), a functional characteristic (the rate of change in the nitrogen content in the environment), and the invariant characteristic (the rate of nitrogen circulation in the system). This has made it possible to clearly classify the data obtained according to completeness, how obtained, how processed, and research goals, which is a cardinal point in determining the mean values of the sets of estimates of variables obtained (extra-system, structural, functional, combined-variants, invariants).

III. Of principle importance is the assumption that when $G = \text{constant} \geq 0$, all estimates of system variables $(\Delta y_i)$ are within the limits of the norm, so that the concept of system "norm" in terms of the norm for the specified properties was concretized, and the region of the norm was determined for the external and internal factors $(\bar{y} \pm \Delta y_i)$, as well as for the disturbing environmental actions $(\bar{x} \pm \Delta x_i)$.

IV. It can be postulated that when $G \to 0$, when the condition $0 < G \neq \text{constant}$ is satisfied, that is, during a period of the decay of homeostasis, all the observed deviations of the individual variables and indices of the specified properties of a system are within the limits of the permissible measure of deviations. This results in the establishment of the "red line" position, designating the boundary prior to the crossing of which the change in system properties caused by a disturbing action is reversible. Deviations beyond the boundary of the "red line" are critical, since when $G = 0$ they will lead the ecosystem to irreversible change in appearance (nonremovable disruptions in properties, pointing to the appearance of a "new" ecosystem) or to its destruction as a system.

The postulates formulated are suitable as initial premises for constructing a numerical model of ecosystem stability, the predictions of which can be readily checked by experiment. One would hope that it can be useful in practical situations arising as a result of the ever-growing effect of man's activity on natural ecological systems.

## BIBLIOGRAPHY

Fedorov, V. D., 1970a, "Special features of the organization of biological systems and the hypothesis of an 'outbreak' by a species in a community," *Vestnik MGU*, Biological Series, No. 2, pp. 71–81.

Fedorov, V. D., 1970b, "The biotic variety of a phytoplankton community and its production characteristics," *Biologicheskiye nauki*, No. 2, pp. 7–15.

Fedorov, V. D., 1972, "The problem of the complex in biology and special features of its resolution," *Vestnik MGU*, Biological Series, No. 6, pp. 30–42.

Fedorov, V. D., Sokolova, S. A., 1972, "The stability of a plankton community and some characteristics of the environment," *Okeanologiya*, No. 6, pp. 1057–1065.

Fedorov, V.D., 1973, "A new index of inhomogeneity in the structure of a community," *Vestnik MGU*, Biological Series, No. 2, pp. 94–96.

Cannon, W. B., *The Wisdom of the Body*, 1932, London.

MacArthur, R. H., "Fluctuation of animal population, and a measure of community stability," *Ecology*, No. 36, 1955, pp. 533–636.

Milsum, H. H., "The nature of living control ssytems," *Engng. J.*, Vol. 47, No. 5, 1964, pp. 34–38.

# Development of Maximum Permissible Environmental Loading—Air

B. J. Steigerwald *

## I. INTRODUCTION

Most of the discussion thus far in the Symposium has focused on the development and selection of air quality goals or standards. Certainly this is a critical decision. In selecting standards we determine the severity of the pollution effect as well as those whom we will protect. Equally important and often more practical are the decisions relative to how we will meet the air quality goal or standard. The consequences of these decisions are often immediate and are very visible and understandable to the public. Essentially, the regulatory strategy or air pollution control scheme that is selected determines the distribution of the burdens of control, and their magnitude.

Clearly, the regulatory strategy that is selected or the maximum permissible environmental loading that is permitted should be the most cost/effective program for the attainment of the air quality goal. In addition we should consider pollution control as additional costs of doing business. Control costs should be paid by the consumer who will benefit from the cleaner air. For example, a community whose control regulations force location of power plants into rural or distant regions should pay a high enough rate for electricity to offset any environmental effects in the area where the plant is located. These may include air pollution, water pollution or effects of strip mining.

Although the challenge is simple and clear—minimize the ratio of costs to effectiveness and distribute burdens fairly—there are many practical constraints in the development of the optimum regulatory strategy. These limitations usually dominate the situation and often preclude the full use of the available analytical tools. Compromises must always be made to account for lack of data on air quality, meteorology, location of receptors and emissions. Political factors also must be considered and these are not necessarily compatible with technical/economic optimization. Despite these limitations the art or science of selecting a regulatory strategy is probably on a par with our ability to set optimum environmental goals. Also, it must always be kept in mind that air pollution is only one small part of a much broader system which should be optimized. Other considerations should include energy supply, transportation, public health care, general economic well-being, recreation, and environmental factors such as garbage disposal and water pollution. Most often these factors cannot be included quantitatively in decision models of air pollution regulatory strategy. Often major compromises must be made with limited quantitative analysis.

## II. FACTORS TO BE CONSIDERED IN ANALYSIS OF PREFERRED REGULATORY PATHWAY

For purposes of the current discussion, we will limit ourselves to the rather narrow problem of developing the best possible air pollution control scheme or set of regulations to meet a well stated air quality goal or standard. Even within this narrow perspective, there are a wide variety of factors that must be considered and analyzed.

### Institutional Factors

Organization of government and the availability of resources often direct the general form of the control strategy or limit the scope of the analyses. These include: form of government-socialistic or capital-include: form of government—socialistic or capital-strong local control programs; statutory authority—freedom provided in enabling legislation, mandatory regulations or exemptions in basic statute; technical sophistication of the nation or local governments developing regulations; resources available for surveillance and enforcement; and extent of public support for air pollution control programs.

### Nature of Pollutants

The effect of the pollutant being considered and the severity of the effect must influence the nature of the control program. The effects spectrum that must

---

*Director, Office of Air Quality, U.S. Environmental Protection Agency be considered is continuous. The spectrum starts at public health and ranges through material damages to esthetics and on to nuisance effects such as visibility and odors.

Where the potential effect is on health, affects a large segment of the population, and is serious, irreversible, and acute, a stringent control strategy must be applied. This would include both the severity of the control measures and their time of application. Examples are beryllium, asbestos, and mercury, three pollutants recently classified as "hazardous" in the United States. Source control for hazardous pollutants is under the very restrictive Section 112 of the Clean Air Act. Section 112 dictates nationwide emission standards, compliance from 90 days to two years, and extensive monitoring and surveillance.

Pollutants such as fluoride whose effect is serious, but not acute and which may not affect large segments of the exposed population, may allow for slightly less priority on the time of application of control measures. Therefore, more reasonable, longer term control options may be possible.

Some effects may call for no control. Recently in the United States, we determined that odor control would not be a nationally manated effort. Odor is deemed a local nuisance and control should proceed at a rate determined by local needs and desires.

## Selecting an Air Quality Strategy

The scientific character of the pollutant is possibly the most critical factor in determining the nature of the control strategy. If there are multiple exposure routes, air pollution control strategy must be integrated into the broader effort and put into perspective with timing and intensity.

Many pollutants such as CO, $SO_2$ and total particulates are emitted directly. Their source/receptor relationship tends to be uncomplicated; air concentrations can usually be related to sources in the same urban area. Several pollutants that are becoming very important, however, are formed in the atmosphere often after long transport distances. Atmosperic chemistry complicates greatly the development of an effective, equitable control strategy. Oxidants, for example, are formed by the reaction in the atmosphere of certain organic compounds, oxides of nitrogen and sunlight. The reaction is complex and can occur within minutes or over several days. Recent data show that this is not strictly an urban problem. Concentrations of oxidant greatly exceeding the air quality standard of 0.08 ppm one-hour average are occurring in many rural areas of the United States 100 or more miles from major cities. Reports include 0.15 ppm in Ft. McHenry, Maryland, which is located 120 miles from Washington, D.C.; 0.13 ppm in Kane, Pennsylvania, which is over 150 miles from Cleveland, Ohio; and 0.5 ppm in Indio, California, located over 120 miles from Los Angeles. Sulfate particulate matter, including sulfuric acid, also seems to be formed in the atmosphere as the result of the oxidation of $SO_2$ gas; only 2% is particulate sulfate. The situation is very complex. Current information in this case is not yet adequate to allow the selection of a workable national control scheme for meeting a sulfate air quality standard.

The sources of a pollutant, their distribution and future growth patterns are also key determinants of the control strategy. To effect significant control, large industrial sources generally install control equipment. Area sources such as home heating or automobiles often require basic changes in fuel supplies, land use patterns or life style of the population. In dry, windy climates such as the Southwestern United States, wind-blown dust, including that from farming and unpaved roads, can be the major source of particulate matter and may preclude meeting the air quality goals with any air pollution strategy.

A formal analysis of alternative control strategies that incorporates many of these considerations recently has been done in the United States for $NO_2$.(1) The analysis resulted in a significant change in the national $NO_2$ control program. A short summary of the study might be illustrative of the value of a preferred control strategy analysis.

After some initial confusion related to the selection of measurement methods, we determined that existing levels of $NO_2$ were not excessive in most U.S. cities. The primary objective of the $NO_2$ analysis was to determine an equitable and efficient air pollution control strategy that would maintain the annual $NO_2$ standard of 100 $\mu g/m^3$ despite the rapid growth of $NO_x$ sources. A proportional model was used to simulate air quality data at five-year intervals out to 1990. The model used current air quality data, six categories of $NO_x$ sources (light, medium, and heavy duty vehicles, industrial processes, area sources, and electric power plants), five sets of growth rates for each source category, two levels of stationary source control, and five $NO_x$ automotive emission standards. A proportional relationship was assumed between $NO_x$ emissions and $NO_2$ concentrations. Annual air quality estimates and control costs were calculated for each strategy. Fuel penalties were also calculated for each level of mobile source control.*

The $NO_x$ analysis summary report recommended that the 1977 automotive standard of 0.4 grams/mile be revised to 2.0 grams/mile. In addition the report recommended that more stringent emission control be placed on new and existing stationary sources in regions with high $NO_2$ concentrations. The recommendation was made after evaluating ranges of air quality, and costs that included fuel penalties. As a result of the analysis, adjustments are now being made to the national control program for $NO_2$ in the United States. Table 1 illustrates an example comparison of two pollution control strategies for New York City. Several other strategies not shown were also compared.

---

*See Attachments: Attachment 1 describes $NO_x$ Analysis; Attachment 2 illustrates a range of estimated air quality and control costs.

## III. GENERAL TYPES OF REGULATORY STRATEGIES

There are two general approaches to developing an air pollution control strategy. The first involves simply the widespread application of reasonably available control technology. The second involves an air quality management program, usually over a relatively small area such as a city, to balance emission control with air quality needs.

### Control Technology

The control technology approach is the more traditional strategy. A general problem is identified and available emission control technology is applied to control the problem. There is little attempt to vary regulations to match air quality levels or to develop long-range plans to maintain the air quality standard. Formal analysis of cost-effectiveness and cost-benefit of alternative strategies is not emphasized in this approach. However, control is technically simple and enforceable. Sources generally know unambiguously their control requirements and can have confidence that they will not change continuously. The initial successes achieved in St. Louis, Pittsburgh and London were based on this form of control strategy.

In the Clean Air Act of 1970, the U.S. Congress generally limited the control technology approach to new sources. Section 111 of the Act, New Source Performance Standards, requires all new major sources of pollution to meet national emission standards set by the U.S. Environmental Protection Agency. These standards are based on application of best demonstrated control technology and take into account the cost of such control systems. National emission standards are set for the lead content of gasoline. Standards are also set for new automobiles, trucks, and aircraft. Standards are also established for selected sources of hazardous pollutants. These are based either on best technology or on the degree of control needed in the most critical situation in the country.**

The control technology approach also is being used increasingly as a compromise solution when air quality management programs cannot be upheld because of lack of data or inadequate analytical tools. Sulfur oxides and particulates are examples, especially in the vicinity of large isolated sources such as power plants in very mountainous terrain where diffusion calculations are unreliable. The control of oxidant in rural locations is another example.

### Air Quality Management

Theoretically, the benefits of the air quality management (AQM) approach are great. Control is required only where needed and only to the extent needed. Usually these benefits cannot be fully realized because of practical limitations, but even if the AQM approach is eventually discarded as the basis for regulations, the AQM analysis will provide valuable experience in setting enforcement priorities and in directing research and data collection programs.

Air Quality Management is required by the U.S. Clean Air Act as the principal approach for the attainment and maintenance of the national air quality standards. The entire country has been divided into 247 Air Quality Control Regions (AQCR). Control region designations are based on considerations such as meteorology, population density and political boundaries. Each state must submit a State Implementation Plan (SIP) that outlines and justifies a regulatory scheme to attain and maintain all of the air quality standards in every region or part of Region within the state. The SIP's are very complex,

**Table 1**
**Comparing $NO_x$ Control Strategies (New York)**

| Strategy | Annual Air Quality ($\mu g/m^3$) | | | | Annual Control Cost ($10^6$ dollars) | | | | Annual Fuel Penalty ($10^6$ dollars) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1977 | 1980 | 1985 | 1990 | 1977 | 1980 | 1985 | 1990 | 1977 | 1980 | 1985 | 1990 |
| Automotive standard of 0.4 grams/mile & Moderate Stationary Source Control | 90 | 90 | 86 | 96 | 144.8 | 276.0 | 457.0 | 510.1 | 67.6 | 134.9 | 212.2 | 227.9 |
| Automotive standard of 2.0 grams/mile and Maximum Stationary Source Control | 91 | 78 | 77 | 87 | 128.7 | 194.7 | 320.6 | 377.4 | 60.5 | 92.8 | 132.1 | 136.7 |

---

**See Attachments: Attachment 3 lists New Source Performance Standard categories and pollutants; Attachment 4 lists dates for controlling Mobile Source Emissions; Attachment 5 illustrates Pollutant and Sources of Hazardous Emission Control.

but each must include information on current air quality, an emission inventory, regulations that will reduce current emissions enough to achieve the air quality standard and evidence of enforcement capability.

The development of a complete AQM program demands great technical sophistication, good understanding of the pollutant and much data. A full discussion of the AQM is beyond the scope of this paper. It seems appropriate, however, to delineate some of the data needed for a full analysis and to describe the status of AQM techniques in the United States.

Information on air quality is vital to the AQM approach. The data must be temporally and spatially compatible with the air quality standard, i.e., hourly, daily or annual measurements and available for key receptor sites and in locations of expected maximum concentrations. Data must be available for enough years to ensure that severe conditions of meteorology and emissions are included. Natural background (wind-blown dust, stratospheric ozone, organics from vegetation) contributions to urban levels must also be determined and included in the analysis. Finally, since the AQM is so quantitative and so dependent on air quality, standard methods of collection and analysis must be developed with known precision and accuracy and the size and siting of monitoring networks must be carefully defined. In the United States there are approximately 4000 air sampling stations, most using standard or equivalent methods and operated by the State or local governments. The data are sent four times each year to a national data system called SAROAD (Storage and Retrieval of Air Data) for organization and evaluation.

Source emission data are the next vital part of the air quality management analysis. The emission inventory should include the magnitude and location of all important sources of the pollutant being evaluated. Large sources should be specified individually along with information on stack height, existing controls, plant age and economic viability. Area sources such as home heating or automobile emissions should be disaggregated into geographical areas compatible with the diffusion model being used. In the United States, this may range from several square kilometers when sophisticated diffusion models are being used to full counties (from a thousand to 30,000 square kilometers) where the simple proportional relationship between emissions and air quality is used. In the United States, approximately seventy percent of the 100,000 large point sources (over 100 tons of emissions per year) and all of the 3300 counties (as area sources) are included in the National Emission Data System (NEDS).

Data on the type and distribution of receptors that may be affected by air pollution may also be necessary in some types of air quality management analysis. The need to include receptors depends upon the type of air quality standard that has been set. Currently in the United States the health related standard must be met in all geographical areas by a date set in the basic legislation (1975 or 1977). Legally, the air quality standard basis for the control regulations and the latest time of application are uniform throughout the country. In practice, there is a priority set on analysis, monitoring, surveillance, and enforcement that favors urban areas or other areas where sensitive receptors are concentrated. Even if allowed by the Act, there are no usable damage functions that relate the magnitude of health effects to changes in concentrations. Damage functions, therefore, cannot be included in any air quality management models.

An optimal control strategy cannot be selected without detailed information on control systems. For each major source category, information for making air pollution control decisions must include effectiveness, availability (including timing) cost and social impact. Data is fairly simple to collect for control systems that involve stack gas cleaning equipment such as precipitators or filters. When pollution control is achieved through a change in the manufacturing process or fuel substitution, we find more difficulty in assigning a cost to pollution control. Some of the more recent control programs involve changes to the transportation system in a city or land use planning which have a social impact that cannot always be stated in simple economic terms. Cost comparisons with more traditional pollution control equipment may not be possible.

The heart of an AQM analysis is the source/receptor model that relates emissions to air quality. In its simplest form this may be a proportional rollback model that assumes, for example, that a 40% reduction in total $SO_2$ emissions for the city will result in a 40% decrease in air concentrations. Unlike more sophisticated diffusion models, the proportional model does not account for source distribution, height of emission or meterological conditions. (2)

In general, diffusion models take into consideration the pollution diluting capabilities of the environment (see Table 2). The more dilution that takes place because of wind, stack heights, or distance between sources, the lower the ambient air concentrations become. The "box" model, "gaussian plum" models, and the CO model are examples of more advanced diffusion models. They allow one to predict the spatial distribution of various concentrations of pollutants for a given day, or over some long time period, which is not possible by proportional modeling.

A current limitation in meterological diffusion models is an inability to include the complex atmospheric reactions that both form and destroy pollutants. This is most serious for photochemical systems involving hydrocarbon, oxides of nitrogen and oxidants. Extensive research is underway and an

## Table 2
### Diffusion Models

| Box Model | Caussian Model | CO Model |
|---|---|---|
| $C_g=\frac{Q}{uhd}$ | $C_g=\frac{Q}{(2\pi)^{1/2}\nu_y hu}\exp\left(-y^2/2\sigma_y V_y^2\right)$ | $C_w=C_r\exp[(0.55)(1-z/z_r)]$<br>$C_l=C_r\exp[(0.87+0.98u_r)(1-z/z_r)]$ |

$C_g$=ground level concentration, $C_r$=roof top concentration, $C_l$=leeward side of building concentration
$C_w$=wind side of building concentration (g $m^{-1}$)
Q =source emissions (g $sec^{-1}$)
u =average wind speed, $u_r$= roof top wind speed (m $sec^{-1}$)
h =mixing height (m)
d =½ [urban area]$^{1/2}$(m)
y =cross wind distance from plume axis (m)
$\sigma_y$ =cross wind standard deviation of plume concentration
$z_r$ =average roof top height (m)
z =receptor height (m)

experimental model has been developed for Los Angeles (3). However, routine use of models that include atmospheric transformations is still several years away and, when available, the need for very detailed data on emissions, ultraviolet radiation and air quality for validation will delay widespread application.

## Management Systems and Models

Two management tools currently used in our AQM programs are particularly good vehicles for identifying and solving problems. Both models help us to determine the effectiveness of control strategies.

The first is the Plan Revision Management System that was recently developed to identify the strengths and weaknesses of the emission regulations found in the State Implementation Plans.*** The Plan Revision Management System uses a computer program to compare the measured air quality values at each monitoring site within an Air Quality Control Region with the estimated air quality values that are computed from the SIP regulations, growth factors, transportation control plans and Federal motor vehicle control programs.

The Plan Revision Management System indicates places where measured levels of air pollution exceed projected levels (see Figure 1). Statistical tolerance intervals are computed from data collected in our data banks and help to indicate the extent of a control strategy's failure to meet the national standards. The management system provides a method for identifying and responding to problem regions throughout the country. It provides a priority ranking for those areas which should receive the most intense review by air quality managers.

The second model is the most comprehensive Air Quality Management tool in the United States. It is called the Implementation Planning Program (IPP) model. (4) IPP incorporates air quality, emissions, meteorological dispersion and control costs for analysis of alternative strategies.

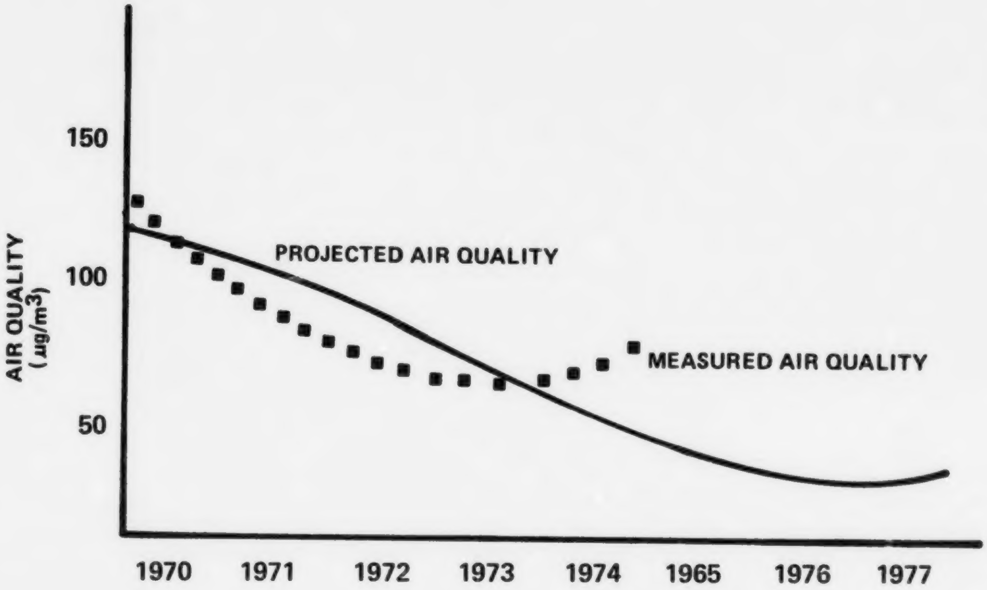
The IPP model is a complex computer program that uses a number of sub-programs (modules) to estimate annual average air quality for $SO_2$ and total suspended particulate. To operate the model, emission data and control cost data are collected from every large source and classes of smaller sources throughout a region. Meteorological data that are also collected include hourly observations of wind speeds, mixing heights, temperatures, and pressures. With these data, the model calculates pollutant contributions from individual sources. Monitored data is used, within one of the modules, to verify the results. Control costs are computed for different emission levels at individual sources. In the final step, individual source pollution contributions and control costs are combined to estimate the cost of control and the ambient air quality associated with implementation of a regional control strategy.

The IPP was used during 1971–72 to model the cities of New York, Buffalo, Philadelphia, and Boston. The model was used to primarily evaluate the effectiveness of the State Implementation Plans. Unfortunately, the massive amounts of data that are required to run the model prohibit more extensive use. The data that is initially put into the model is specified by a rigid format that requires a complete emission inventory to be matched with meteorological data and the air quality monitoring data.

IPP is constantly being improved. New computer programming techniques are being developed to make the model more efficient. A sub-program is being developed to estimate short-term pollution concentrations as well as annual concentrations. (5, 6) In addition, IPP is being reprogrammed to update the pre-set control technique that reflects current control costs and technology (see Table 3).

---

***Attachment 6 illustrates control curves for the Plan Revision Management System.



CALENDAR YEAR

FIGURE 1. PLAN REVISION MANAGEMENT SYSTEM

## IV. MAINTENANCE OF AIR QUALITY STANDARDS

Programs to maintain air quality standards generally allow for much more freedom of action and imaginative solutions than control schemes to rollback pollution to attain the standards. Since the emphasis is on new sources, evaluations can be made of the effects of city planning, new mass transit systems, effective solid waste disposal, industrial siting practices, economic incentives and emerging control technology. Theoretically, programs to maintain standards should accommodate themselves fully to balancing air quality with economic and population growth. However, there is little experience with such programs and no working analysis techniques that have widespread applicability. Since these are future programs, there is still time for research to make significant contributions.

Efforts to maintain air quality in the United States are a complex program of activities, many of them required by the Clean Air Act. Current state emission regulations must be restrictive enough to achieve the standards in the most polluted areas; therefore, where regulations apply state-wide, there will be excess control or "overkill" in all other areas. This "overkill" will allow much growth of new sources to be absorbed without violation of standards. For example, $SO_2$ controls needed in Philadelphia will reduce $SO_2$ levels in surrounding counties by about 40% (15 to 25 $\mu g/m^3$) even though they are now well below the standard. Special emission regulations for new sources are a major part of programs to maintain the standard. The Clean Air Act and EPA regulations require nationwide emission standards for many new sources including automobiles and trucks, gasoline additives, aircraft and all major industrial processes. These help ensure minimum air quality impact of future economic and industrial growth. In addition all new industrial and combustion sources must obtain a pre-construction approval which can be given only if it is determined that the air quality standard will not be violated in the vicinity of the plant. Recently these regulations have been expanded to include a pre-construction review of airports, highways, shopping centers, amusement centers, and other "indirect" sources that attract enough automobile traffic to threaten the air quality standard. Finally, a comprehensive analysis is being made of needed maintenance programs for the more populous Air Quality Control Regions in the country. These include projections to 1985 for emissions and air quality and development of appropriate regulations to ensure maintenance of the standards, including consideration of land use and transportation.

## Table 3
## Implementation Planning Program (IPP)

PURPOSES:
- (a) Relate emissions of sulfur dioxide and particulate matter to ambient annual concentrations of the two pollutants
- (b) Test regional control strategies of cost-effectiveness in reducing ambient pollutant concentrations

PROGRAM MODULES

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| *Source Data Management Segment* | *Air Pollution Diffusion Segment* | *Control Cost Segment* | *Emission Standards Segment* | *Regional Strategies Segment* |
| Purpose: Create and maintain emissions inventory data for use with other modules | Purpose: Calculate pollutant contributions from individual sources to preselected receptors | Purpose: Calculate emission reductions & costs associated with applications of various control techniques | Purpose: Determine for each source the most cost-effective control technique needed to achieve input emission stds. | Purpose: Estimate cost of control and ambient air quality associated with implementation of a regional control strategy |

FUTURE IMPROVEMENTS

1) Update pre-set control technique information to reflect current control cost and technology
2) Include, in the Air Pollution Diffusion Segment, the Samples Chronological Input Model which has the capability of calculating short term concentrations

## V. SUMMARY

This paper has attempted to present some of the conceptual considerations that we feel are important in developing optimum control strategies to achieve and maintain air quality standards. Strategy development is a complex, very practical, and vital part of the total field of air pollution control. The control strategy and resulting regulations determine not only who will control but also the severity of the necessary control measures. No attempt has been made to discuss the detailed techniques of air quality management analysis. Entire books have been written on many of them.

Currently in the United States, the concepts of air quality management are used primarily to guide the development of control strategies. A full quantitative analysis is seldom the determining factor in setting regulations because of institutional and political constraints and data limitations. Despite many of the unquantifiable factors associated with air quality management, improvements in models such as IPP and development of new techniques similar to the Plan Review Management System will possibly provide more objectivity in the selection of future control strategies.

## BIBLIOGRAPHY

1. $NO_x$ Analysis Report, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, September 21, 1973.
2. Proceedings of Symposium on Multiple-Source Urban Diffusion Models, U.S. Environmental Protection Agency, 1970.
3. Photochemical Smog, Proceedings of International Symposium on Air Pollution, October 17, 1972, Tokyo.
4. Air Quality Implementation Planning Program, Volume 1, National Technical Information System—PB. 198299.
5. A Users Manual for the Sampled Chronological Input Model, GEOMET, Inc., Contract No. 68–02–0281, 1973.
6. Validation and Sensitivity Analysis of the Gaussian Plume Multiple-Source Urban Diffusion Model, GEOMET, Inc., Contract No. CPA 70–94, November 1971.

## Attachment 1
## Description of $NO_2$ Analysis

- COMPUTER MODEL
- EMISSIONS OF $NO_x$ AND CONCENTRATIONS OF $NO_2$
- SOURCES OF $NO_x$ EMISSIONS
  - Light Duty Vehicles
  - Medium Duty Vehicles
  - Heavy Duty Vehicles
  - Electric Power Plants
  - Industrial Plants
  - Area Sources
- GROWTH FACTORS
- AUTOMOBILE EMISSION STANDARDS
- STATIONARY SOURCE CONTROL
- COSTS

## Attachment 2
## Ranges of $NO_2$ Air Quality Projections and Costs

Air Quality ($\mu g/m^3$)

Costs (Millions $)

| City | 1975 | 1980 | 1985 | 1990 |
|---|---|---|---|---|
| Phoenix | 80 | 58–71 | 56–78 | 61–80 |
| | 9.4–9.6 | 16.5–39.3 | 26.0–69.9 | 30.3–85.0 |
| Los Angeles | 181 | 97–132 | 94–153 | 104–174 |
| | 64.8–66.1 | 103.4–244.3 | 136.9–396.0 | 142.3–447.7 |
| San Francisco | 86 | 67–75 | 68–87 | 80–101 |
| | 28.6–29.2 | 45.6–108.8 | 62.4–177.5 | 65.2–209.9 |
| Denver | 78 | 72–86 | 81–101 | 95–120 |
| | 8.4–8.6 | 13.4–45.0 | 18.0–68.2 | 18.8–79.0 |
| New York | 101 | 74–97 | 68–99 | 74–112 |
| | 75.3–76.8 | 119.0–274.2 | 168.1–478.6 | 176.4–562.3 |
| Philadelphia | 88 | 74–97 | 82–109 | 95–127 |
| | 30.2–30.8 | 47.2–113.0 | 61.8–183.2 | 63.1–208.7 |
| Washington, D.C. | 94 | 69–103 | 64–102 | 72–120 |
| | 15.5–15.8 | 24.2–60.1 | 34.6–103.4 | 37.0–121.5 |
| Chicago | 120 | 97–117 | 93–127 | 103–146 |
| | 38.6–39.4 | 58.7–148.0 | 83.4–245.2 | 87.1–279.5 |
| Baltimore | 107 | 84–102 | 78–107 | 89–132 |
| | 12.3–12.6 | 22.7–57.4 | 39.6–118.5 | 50.9–164.9 |
| Salt Lake City | 64 | 54–60 | 45–60 | 48–67 |
| | 5.4–5.5 | 9.6–22.4 | 14.5–40.1 | 16.8–49.5 |

## Attachment 3

### New Source Performance Standard Categories and Pollutants

| | |
|---|---|
| Steam Generators | $SO_2$, $NO_x$, TSP |
| Municipal Incinerators | TSP |
| Portland Cement Plants | TSP |
| Nitric Acid Plants | $NO_x$ |
| Sulfuric Acid Plants | $SO_2$ |
| Asphalt Concrete Plants | TSP |
| Petroleum Refineries | $SO_2$, HC |
| Secondary Lead Smelters & Refineries | TSP |
| Petroleum Storage | HC |
| Secondary Brass & Bronze Refining | TSP |
| Iron and Steel Mills | TSP |
| Sewage Treatment Plants | TSP |

## Attachment 4

### Dates for Controlling Mobile Source Emissions

| | Automobiles | Light Trucks | Heavy Trucks | Aircraft | Motorcycles |
|---|---|---|---|---|---|
| $NO_x$ | 1976 | 1975 | 1974 | 1981 | 1976 |
| HC | 1975 | 1975 | 1974 | 1981 | 1976 |
| CO | 1975 | 1975 | 1974 | 1981 | 1976 |
| Particulates* | NA | NA | NA | 1979 | NA |

*NA=Not Available

## Attachment 5

### Hazardous Emission Control—1973

| Beryllium Plants | Asbestos | Mercury |
|---|---|---|
| Extraction Plants | Mining & Milling | Mercury Ore Processing Facilities |
| Foundries | Manufacturing | |
| Ceramic Manufacturing Plants | Fabrication<br>Demolition | Mercury Cell Chlor-Alkali Plants |
| Machine Shops | Spraying | |

# Estimating the Maximum Permissible Intensity of the Complex Effect of Chemical Factors in the Production, Municipal, and Domestic Environment on Man

N. F. Izmerov *

Investigations of the harmful effect of chemical compounds on living organisms when the toxins enter primarily from some one environment (air or water, food, and so forth) have now taken shape. Approaches to the hygienic limitation of pollution of individual objects in the habitat have been developed in this plan. At the same time, we know that the same harmful substances can enter the body simultaneously from the air in production areas and in living quarters, from the atmosphere of populated areas, from water and food, from clothing and shoes, and so forth.

And it has been established that emissions from the production of aluminum, cryolite, and superphosphates settling on soil and water surfaces create artificial geochemical provinces with increased transfer of fluorine to plants and animals ecologically associated with man. There are individual microregions 1.5 to 2 km from a cryolite plant in which the fluorine content in the air was found to measure 0.46–0.85 mg/m$^3$, exceeding the maximum permissible concentration (MPC) for air by a factor of 40–80. The fluorine content in the soil 4 km from the plant exceeded the natural content by a factor of 1.5–27.5. The harmful effect of plant emissions on trees, as well as on the health of animals and man, has been noted (Sadilova, M.S., 1964). The results of an investigation of individual groups in the population living around aluminum and cement-slate plants where elevated fluorine content is present in various environmental objects revealed fluorosis in varying degrees of seriousness in children and in the adult population (Audera, A. K., 1966), as well as change in the thyroid gland resembling enthyroid goiter (Kharitonova, V. A., Perfil'yeva, Z. V., 1966). Natural investigations of animals (Lindberg, Z. Ya., 1965) confirmed the prevailing role of production emissions in the formation of similar geochemical provinces with increased fluorine content in environmental objects.

The population in agricultural areas absorbs a big part of the pesticides in the course of everyday living and not from the air in the work area, while working in agriculture. The specific content of DDT in food products varies from 92 to 95%, in water 5%, and in air 3%, of the total dose absorbed. The effect of pesticides in a work area can be characterized by its transitory nature, and by its periodicity (L. I. Medved', et al., 1971). The complex nature of the effect on man is characteristic of a great many chemical compounds. One of the most widely propagated production toxins, for example, is carbon monoxide, yet is equally a domestic toxin. The products of the incomplete combustion of natural and artificial gas, and of other types of fuel in buildings, vehicle exhaust gases, and emissions from municipal and domestic enterprises, are the primary sources of air pollution, along with emissions from metallurgical industry plants, and the like. Signs of the effect of carbon monoxide (increased blood level of carboxyhemoglobin) have been detected in 78% of those examined and living in apartments supplied with gas in individual microregions (Martynyuk, V. Z., Devyatko, D. G., 1956).

Hydrocarbons affect the inhabitants of oil drilling areas during production activities, as well as when production is not taking place. Changes in the cardiovascular system (lability of arterial pressure, changes in the vegetative regulation of vascular tonus) have been found in workers in refineries refining high-sulfur crude oil, as well as among the population living in the surrounding area (Zagidullin, Z. Sh., 1966).

However, it should be pointed out that some substances (carbon disulfide, benzene, carbon tetrachloride, and others, in particular) are primarily industrial toxins, whereas others, natural food and fungous toxins, for example, are primarily domestic toxins.

A toxicological estimate of the complex effect of toxins on the body (upon simultaneous intake from

---

*Director, Order of the Red Banner of Labor Scientific-Research Institute of Industrial Hygiene and Occupational Diseases of the Academy of Medical Sciences of the USSR different environmental objects) cannot be identified with the combined effect of several toxins, and presents very definite difficulties.

The thrust of the attention given to resolving questions of setting complex hygienic norms for chemical compounds naturally is concerned with threshold and harmless levels of toxin effect on the organism upon isolated admission. Experimental research in particular has shown that when there is simultaneous admission of chemical substances (12 nonelectrolytes) through the gastrointestinal tract and the respiratory organs at the MPC level or at the threshold of chronic effect, there is summation of the effects (S. M. Pavlenko, 1973). Ye. I. Spynu (1967, 1968) has calculated that the summed quantity of pesticides entering the body from different environments at the MPC level exceeds the harmless dose. He therefore proposes establishing a harmless maximum permissible dose of a drug for man, obtained by totaling the maximum permissible doses and concentrations in all environments, with consideration given to their specific importance:

$$D_m = \sum_{i=1}^{4} D$$

where

$D_m$ is the maximum permissible harmless dose;
$D_1$ is the dose of pesticide in a food area ;
$D_2$ is the dose of pesticide entering with the inspired air in the work area;
$D_3$ is the dose of pesticide entering with the water;
$D_4$ is the dose of pesticide entering with the inspired atmospheric air.

Inherent in the realization of his proposals is the difficulty that arises because of the need to determine the specific importance of the MPC in the different environmental objects in order to calculate the maximum harmless dose.

Also proposed for calculating the maximum harmless dose for simultaneous admission of a single substance into the body from different environments, is the use of the principle of setting hygienic norms as applicable to the combined effect of several chemical compounds, the effect of which is unidirectional in nature and based on the additive effect (A. I. Korbakova et al., 1971). The use of a known formula for the summed effect has been suggested for these purposes:

$$\frac{\text{From prod.}}{\text{MPC for prod}} + \frac{\text{From atmosphere}}{\text{MPC for atmosphere}} + \frac{\text{From water}}{\text{MPC for water}} + \frac{\text{Dose from food}}{\text{POC for food}} \leqslant 1$$

The availability of norms for many pesticides, and for certain other substances, in the air in work areas, in the atmosphere, in reservoirs, and in food products, as well as the use of formulas for calculating approximate levels, today enables us to obtain some idea as to the summed content of substances in the environments, and to recommend permissible quantities in each of them. At the same time, the differences in the approaches taken to justify the MPC in different environments make it difficult to use the principle described. We know, for example, that only the MPC in the air in work areas and the mean daily MPC in the air in populated areas can be established from the resorptive effect (the maximum single concentration can be established from the reflex effect), and sometimes in terms of the toxicological criterion for reservoir water for sanitary and domestic purposes, and for food (with a very small safety factor = 100). This behavior pattern can be, and is, the most frequent dependency, but it fails to take into consideration the possibility of a different effect for the different paths over which admission can occur. Toxicity, as well as the danger of poisoning, depends on how certain chemical compounds enter the body. Fluorodinitriles, for example, are only slightly toxic when injected into the stomach ($LD_{50}$ = 997–2917 mg/kg) (Korshunov, Yu. N., 1969), but are highly toxic when inhaled and the danger of developing acute and chronic poisoning is very high (these substances, when inhaled, are included among the highly toxic and dangerous compounds) ($LD_{50}$ = 62–67 $mg/m^3$, $Lim_{ch}$ = 1 $mg/m^3$).

It would appear that aliphatic nitriles injected into the stomach undergo enzymatic conversion in the liver and become less toxic by a factor of 100. The neutralization of fluorodinitriles in the acirc environment of the stomach also is of definite importance.

The different nature of the effect stimulating toxins have on the body when they enter over different paths is important. Thus, bromine is a specific stimulating toxin when inhaled, and the changes in the respiratory system are of decisive importance at all levels of its effects, and for any duration (Ivanov, N. G. et al., 1973). But when bromine enters the body over other pathways, changes in the central nervous system, in the reticular formation, in the cerebral cortex (Krylov, O. A., 1960; Z. Ya. Dolgova, Ye. G. Dolgov, 1964), and in the thyroid gland (Verkhovskaya, N. N., 1962) come to the fore. Here the problem of fixing a complex norm is much more complicated.

We also must pause to consider the important differences that exist between the actions of chemical compounds that take place under natural conditions upon inhalation, and when they enter via the gastrointestinal tract. In the case of inhalation we most often must deal with the intermitten nature of the effect, whereas when admission is over an enteral pathway the mode usually is monotonic. It now has been established that the intermittent mode of action in some cases is more harmful to the body. The intermittent action of such nonelectrolytes as toluene (Babanov, G. P. et al., 1972), carbon tetrachloride, methylene chloride (I. P. Ulanova, et al., 1972), trichloroethylene (Lomonova, T. V., 1973), and a number of pesticides (Burkatskaya, S. N., Mityushina, V. I., 1968) resulted in the development of chronic intoxication and failure of the adaptive mechanisms more rapidly than their monotonic action. The hygienic significance of the variations in concentrations of harmful substances is different for different toxins, and is determined by their properties. In particular, the solubility of substances in water and in fats, the rate of saturation of the blood by a substance, and of its conversion in the body, the pathways and rate of excretion from the body, and the ability to accumulate, all play a significant role.

Moreover, long interruptions between the effects of most chemical substances can result in a reduction in the toxic effect (Sidorenko, G. I., Pinigin, M. A., 1971). At the same time, the biological effect can be more pronounced when interruptions are brief (N. A. Tolokontsev, 1960; Lyublina, 1973). These facts also serve to definitely complicate the problem of establishing complex norms for chemical compounds.

Data that show the practical necessity for an integrated approach to establishing hygienic norms for compounds that are all present at the same time in several of man's habitats are now at hand. Experiments proving, in particular, the inadequacy of the MPC for the fluorine in water, and ion in the atmospheric air when the effect is complex (Petina, A. A., 1967), as well as in the water, and in the air in a work area when the effect is complex (Avilova, G. G., et al., 1973), have been devised.

A comparatively short biological experiment (1 month) showed that in the case of complex entry into the body of the fluorine ion with water and air at concentration at the MPC level for individual environments, the summed dose of fluorine in the body causes changes in many of the biochemical indices of the carbohydrate and mineral metabolisms, as well as in the integral indices of body condition (Avilova, G. G., et al., 1973).

We therefore can conclude that the problem of the complex effect of chemical compounds as a whole is not yet solved. The need for hygienic norms for chemical compounds entering the human body simultaneously from different environmental objects is an urgent one, and requires the development of theoretical and methodological approaches that will provide for detailed resolution of this problem. Today it apparently is still too early to talk about a single norm that will express the integral threshold of the prolonged effect and the integral MPC.

The biological magnitude of an integral MPC, established using highly sensitive enzymological tests to establish exposure, can be an important index of the summed effect of a chemical factor for a number of substances, and this is in addition to the environmental estimate.

The presence of these substances in the environment can be limited by making a determination of the permissible content of the products proper, or of their metabolites, as well as by the state of the enzyme systems in the biological media, or of the most sensitive organs.

It should be pointed out that the establishment of two magnitudes characterizing the maximum permissible quantities of a substance entering through the lungs and through the gastrointestinal tract must today be recognized as the most realistic approach when setting complex norms. At the same time, the MPC for a substance entering the body via the respiratory organs must take into consideration the possibility of its entering the body with the air breathed in a production area and from the atmosphere.

The MPC for a substance entering through the gastrointestinal tract should take into consideration entrance with water and food.

Efforts such as these are particularly important in the case of substances that have far-reaching effects (mutagenic, embryotropic, carcinogenic).

Approval, in the form of legislation, of hygienic norms for 3,4-benzpyrene in individual environments (the air in a work area, atmospheric air, and water) is a considerable step forward in this regard. At the same time, it has been shown that there is little difference between the minimum effective and maximum nonblastomogenic doses of 3,4-benzpyrene for any of the pathways over which entry is had into the body. The same blastomogenic effect can be observed, approximately, in the case of close summed doses of 3,4-benzpyrene injected into animals in different ways (intratracheally, cutaneously, and into the stomach), as well as when the data are converted for man (Yanysheva, 1972; Shabad, L. M., Sanotskiy, I. V., et al., 1973).

Thus, despite the natural difficulties, an integral estimate of the maximum permissible intensity of the complex effect of chemical compounds on man from different environmental objects not only is necessary, it is possible.

## BIBLIOGRAPHY

1. Audera, A. K., *Vliyaniye promyshlennykh vrednostey na sostoyaniye rotovoy polosti* [The influence of harmful industrial conditions on the condition of the oral cavity]. Riga, 1966, pp. 34–37.
2. Babanov, G. P., Burov, Ya. A., Skolbey, N. A., Verkhovskiy, L. G., Abramyan, G. G., Tritskaya, I. L., Isakhanov, T. M., *Toksikologiya i gigiyena produktov neftekhimii i neftikhimicheskikh proizvodstv* [The toxicology and hygiene of petrochemical products and petrochemical production], 1972, pp. 32–45.

3. Burkatskaya, Ye. N., Mityushina, V. I. *Gigiyena primeneniya, toksikologiya pestitsidov i klinika otraleniy* [Hygiene of the use and toxicology of pesticides and the clinical aspects of poisoning], Kiev, 1968, p. 698.
4. Verkhovskaya, I. N., *Brom v zhivotnom organisme i mekhanism yego deystviya* [Bromine in the animal organism and the mechanism of its effect], Moscow, 1962.
5. Zagidullin, Z. Sh., *Trudy Ufimskogo institute gigiyeny truda i profzabolevaniy*, Vol. 16, No. 2, 1965, pp. 22–26.
6. Ivanov, N. G., Germanova, A. L., Pozdnyakov, V. S., Klyachkina, A. M. et al., In Book: *Adaptatsiya i kompensatsiya pri khimicheskikh vozdeystviyakh* [Adaptation and compensation in the case of chemical effects], Moscow, 1973, pp. 75–92.
7. Korbakova, A. I., Shumskaya, N. I., Zayeva, G. N., Nikitenko, T. K., *Nauchnyye osnovy sovrem. metodov gigiyenicheskogo normirovaniya khimicheskikh veshchestv v okruzhayushchey srede* [Scientific bases of modern methods for setting hygienic norms for chemical substances in the environment], Moscow, 1971, pp. 35–40.
8. Korshunov, Yu. N., *Toksikologiya novykh prom. khim. veshchestv*, Moscow, No. II, 1969, pp. 79–85.
9. Krylov, O.A., *Fiziologiya zhurn. im. Sechenova*, Vol. 52, No. 7, 1966, pp. 906–910.
10. Lindberg, Z. Ya., *Materialy 13 nauch. sessii Rizhskogo med. institute* [Materials from the 13th session of the Riga Medical Institute], Riga, 1965, pp. 53–55.
11. Lomonova, G. V., Frolova, I. N., Gritsevskiy, N. A., *Voprosy psikhofiziologii gigiyen. analiza trudovykh protsessov* [Questions of the psychophysiology of a hygienic analysis of labor processes], Moscow, 1967, pp. 57–67.
12. Lyublina, Ye. I., *Kolichestvennaya toksikologiya* [Quantitative toxicology], 1973.
13. Medved', L. I., Spynu, Ye. I., Burvy, V. S., Belonozhko, G. A., Antonovich, Ye. A., Vrochinskiy, K. K., "Scientific bases of modern methods for setting hygienic norms for chemical substances in the environment," *Materialy vsesoyuznoy konf. 21–22 oktyabrya 1970 g.* [Materials from the All-Union Conference of October 21–22, 1970], Moscow, 1971, pp. 13–17.
14. Pavlenko, S. M., *Gigiyena i sanitariya*, No. 1, 1972, pp. 40–44.
15. Avilova, G. G., Golubovich, Ye. Ya., Mal'tseva, N. M., Pankratova, G. N., et al., In book: *Adaptatsiya i kompensatsiya pri khimicheskikh vozdeystviyakh* [Adaptation and compensation in the case of chemical effects], Moscow, 1973, p. 92, 103.
16. Petina, A. A., Yel'nichnykh, L. N., In book: *Flyuoroz i yego profilaktika* [Fluorosis and its prevention], Sverdlovsk, 1967, pp. 50–55.
17. Sadilova, M. S., *Voprosy gigiyeny i profpatologii i toksikologiya* [Questions of hygiene and industrial pathology and toxicology], Sverdlovsk, 1967, pp. 25–26.
18. Sidorenko, G. N., Pinigin, M. A., "Scientific bases of modern methods for setting hygienic norms for chemical substances in the environment," *Materialy vsesoyuzn. konferentsii 21–22 okt. 1970* [Materials from the All-Union Conference of October 21–22, 1970], Moscow, 1971, pp. 13–17.
19. Spynu, Ye. I., *Gigiyena primeneniya, toksikologiya pestitsidov i klinika otravleniy* [Hygiene of the use and toxicology of pesticides and the clinical aspects of poisoning], Kiev, 1968, pp. 103–108.
20. Tolokontsev, N. A., *Gigiyena i sanitariya*, No. 3, 1960, pp. 29–35.
21. Ulanova, I. P., Avilova, G. G., Bazarova, L. A., Mal'tseva, N. M., Migukina, N. V., Khalepo, A. I., Eytingon, A. I., In book: *Adaptatsiya i kompensatsiya organisma pri vozdeystvii khim. zagryuazneniy okruzh. sredy* [Adaptation and compensation by the organism to the effect of chemical pollution of the environment], Moscow, 1972, pp. 17–21.
22. Kharitonova, V. A., Perfil'yeva, Z. V., *Voprosy profilaktiki i lecheniya zoba na Urale* [Questions of the prevention and treatment of goiter in the Urals], a collection of papers from the Sverdlovsk Medical Institute, Sverdlovsk, 1961, No. 34, pp. 20–25.
23. Shabad, L. M., Sanotskiy, I. V., Zayeva, G. N., Bruyevich, T. S., Katsnel'son, B. A., Yanysheva, N. Ya., Shugayev, B. B., *Gigiyena i i sanitariya, Moscow*, No. 4, 1973, pp. 78–81.
24. Yanysheva, N. Ya., *Gigiyena i sanitariya*, No. 7, 1972, p. 87.

# The Concept of the Threshold Nature of the Reaction of Living Systems to External Actions and its Consequences in the Problem of Protecting the Biosphere Against Chemicals

I. V. Sanotskiy *

Legislative and other types of regulations of the state of man's habitat, and that of the plant and animal species associated with him, the toxicological limitation (hygienic normalization) of pollution of the production sphere and of daily life, as well as the derived practical measures of a technological or other nature, depend in the final analysis on the resolution of a fundamental question. Is there a threshold of intensity of *harmful* effect of external factors, or is any intensity different from zero harmful? In other words, can objective scientific facts be used as the basis for protecting the biosphere, or should subjective decisions (the "acceptable risk" concept on the one hand, and so-called "zero" pollution on the other, and so forth) be used to some extent?

What must be emphasized is the fact that one generally is not talking about the *threshold of any changes* in living systems as a result of external excitation, but of the incidence of reactions beyond the limits of ordinary physiological variations observed in the process of homeostasis, something that will be taken up in more detail below.

A mathematical description of the task is not the subject of this report. The purpose of this report is to set forth corresponding points of view, and some of the material used as the basis for for them.

As is known, the occurrence of excitation, that is, of a complex biological reaction to external action in the form of functional changes, and the physicochemical processes that underlie them, is a property of living matter at any stage of organization.

The magnitude of the effect depends on the strength of the stimulation. Bowditch (1871) discovered the "all" or "nothing" law, that is, there is no excitation ("nothing") when stimulations are subthreshold, but maximum excitation ("all") is developed when the stimulus is at threshold strength, and does not increase when the stimulus is intensified. However, these propositions now are regarded only as a rule, because local excitation, local response, occurs in the stimulated sector when stimuli are subthreshold. Nor can the "all" characterize the maximum that the action potential of the nerve or muscle fiber can attain. According to another law, the "law of intensity," the effect of the reaction is greater the stronger the stimulus causing it, over a predetermined range of intensity. Even more complex interrelationships, three and more phases, have been described (Vvedenskiy, N. Ye.; I. V. Sanotskiy, 1950, 1970; P. V. Simonov, 1962). The magnitude of the reaction, as is known, depends significantly on the functional state of the reacting substrate.

## Content of the Problem

The question of the feasibility of establishing a threshold of harmful effect for most types of effect of chemical compounds is not in doubt among most authors. However, this question is debatable when it comes to mutagens, blastomogens, and radiation injuries. No decision was reached at the symposium organized especially for the discussion of this problem in February 1970, by the American Association of Hygienists.

At the same time, the threshold nature of the effect of all harmful factors (including radiation, and particularly chemical) can be proven theoretically by starting from the basic differences between living and nonliving things (the constant exchange by the organism of matter and energy with the environment, the continual restoration by the organism of its structure, and the constant adaptation to the environment, including those similar to itself during the process of reproducing).

The simplified expression

$$D_\gamma = D_0 - D_{ex} - D_{met}$$

where $D_\gamma$ is the dose of toxin reaching the receptor; $D_0$ is the dose of toxin introduced into the organism; $D_{ex}$ and $D_{met}$ are the doses of toxin excreted from the organism, and rendered harmless as the toxin moves toward the receptor, respectively, shows the threshold nature of the effect [provided that ($D_{ex} + D_{met}$) is not decreased in proportion to decrease in $D_0$]. Data from an experiment with nonmetabolized and metabolized substances show that this condition is observed. Indeed, only hypothetical substances absolutely not excreted and absolutely not rendered harmless can reach the receptors without decreasing in quantity. There are no such substances.

---

*Institute of Labour Hygiene and Professional Diseases, Moscow

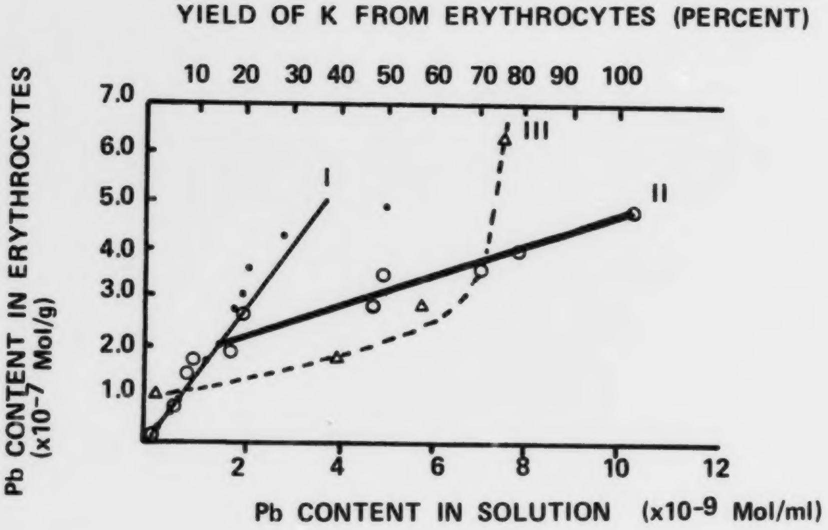
However, even if the receptors are considered to be isolated, so to speak, from the integral organism, it will be found that simple fixation of the toxin by the receptor still does not mean a biologically significant reaction with it. For example, it has been shown in isolated human erythrocytes that fixation (including strong) up to a certain limit by a cell membrane of lead ions does not lead to an increase in the yield to an incubation medium of potassium and hemoglobin (Figure 1), according to data provided by Grigarzik and Passow, 1958.

Thus, the reconstruction of the molecules can take place without any obvious functional changes in the biosubstrate. The occurrence of a "biologically important" effect depends on the "significance" of the substrate in the vital activities of the cell, or on the existence of a reserve metabolic pathway (Dinman, 1972).

The processes of adaptation, and of constant restoration (activation of detoxication, intensified production of biological structures bound by the toxin, such as enzyme molecules, the inclusion of bypasses for normal metabolism, the activation of the organism's macrosystems aimed at adaptation to new living conditions, and others, for example) are added to this in the integral organism. Thus, damage develops only when the rate at which the damaging processes (inactivation, degeneration, and so forth) occur exceeds the rate at which the processes of adaptation and restoration occur. These phenomena were well studied during the fractional use of ionizing radiation, and these are the facts that became the basis for the ideas as to the threshold nature of its effect (Strelin, G. S., et al.).

Similar phenomena have been described as taking place during the action of chemical factors. Thus, cessation of intensified metabolization of bromobenzene in p-bromophenylmercapturic acid (during extended action) was accompanied by decompensated pathology (T. Shamilov). An increase in the expiration of methylene chloride, and a corresponding reduction in its content in the blood, led to physiological adaptation (with a conditionally monotonic 1.5-month effect) (G. G. Avilova and N. M. Mal'tseva).

Adaptation at the subcellular, cellular, organ, organism, and population level while under the extended effect of a damaging factor has repeatedly been demonstrated when the external effect has been most varied in its nature. Included here is the repair of chromosome rearrangements in the case of fractional irradiation (N. P. Dubinin, S. P. Frmonenko,



SOURCE: GRIGARZIK H., PASSOW H., PFLUGERS ARCHIV. VOL. 276, p. 80 (1958)

FIGURE 1. THE Pb CONTENT IN ERYTHROCYTES AND THE YIELD OF K FROM THE ERYTHROCYTES AS A FUNCTION OF THE Pb CONTENT IN SOLUTION

et al.), the normal life of human populations under increased (up to a predetermined limit) radiation background conditions (high mountains, areas in which radioactive ores are present, and the like), repair of chromosome rearrangements in the case of the extended effect of mutagens: ethylenimine, L. D. Katosova, 1972; ethylene chlorohydrine, G. K. Isakova, et al., 1971; 2-methylthio-4,6-bis-(isopropylamino)-sym-triazine (prometrin), DDT, 2,4D, L. P. Yefimenko, 1973, v. N. Fomenko, et al., (Figure 2), rapid inherited adaptation of bacterial to antibiotics, of insect populations to insecticides, and so forth.

The law of natural selection is not applicable to a human population, so the precise facts as to the true adaptation of man to life under chemical pollution conditions have not yet been forthcoming.

The theory of *harmfulness* of the effect is of great importance in the problem of the threshold nature of an organism's reactions to external effects. What is important to the protection of the environment is not any threshold of reaction of the biosubstrate, but the threshold of reactions that have the harmfulness criterion. Similar ideas were formulated by N. S. Pravdin in 1934, and the theory was developed by his successors. More and more scientists have recently subscribed to the propositions expressed. Weil (1972), for example, believes there are dose levels that have no harmful effects on any animal species or on humans, although the effect may be statistically meaningful. I. G. Akoyev, et al. (1972), while confirming the nonthreshold nature of primary reactions to radiation effect, advanced the hypothesis of a threshold number of damaged systems; that is, the hypothesis of a threshold of *harmful* effect.

Thus, the theory of the threshold nature of the effect does not imply reaction of the biobustrate to the effect of external factors in general, but rather reactions that are biologically (including medically) important.

## Analysis of Dose-Effect Curves

The threshold nature of the harmful effect of external factors can be demonstrated by an analysis of the dose-effect curves. Many of the ideas that have evolved as to the nonthreshold nature of the effect have been the result of assumptions and extrapolations. The dose-effect relationship is not linear, so the establishment of the end of the curve corresponding to the lower level is extremely important (Weil, 1972; Stokinger, 1972; Weisburger, 1968; Bock, 1968; Mantee, et al., 1961, 1963; Brues, 1958). The conditions under which the experiment is performed (species, sex, age of animals, state of nutrition, and so forth) have a significant effect on the shape of the curve; so the results of determining thresholds should be reproduced many times over.

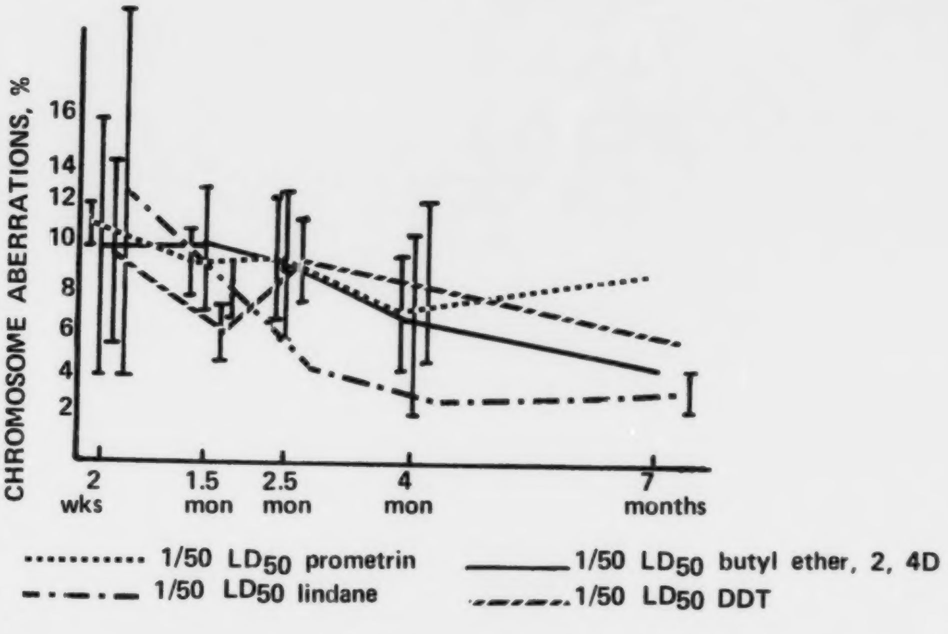


FIGURE 2. THE CYTOGENETIC EFFECT OF PROMETRIN; LINDANE; BUTYL ETHER, 2, 4D; AND DDT IN THE BONE MARROW OF RATS

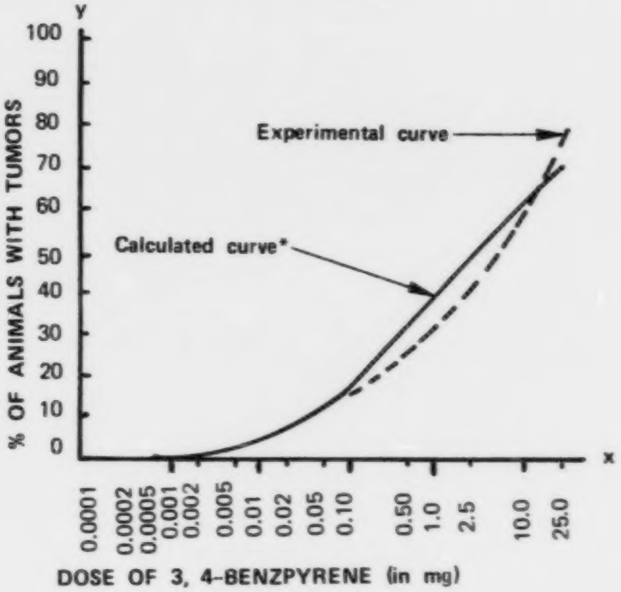
The intensity of the organism's reaction usually decreases with reduction in the dose of the substance administered, with the reaction dropping to zero before the dose reaches zero (Hatch, 1973).

As in the past, most arguments today stem from analysis of dose-effect curves obtained as a result of the effect of mutagens and blastomogens. The curve of the blastomogenic effect of 3,4-benzpyrene obtained by N. Ya. Yanysheva (Figure 3) enabled the USSR to be the first in the world to establish the maximum permissible concentration for this toxin (L. M. Shabad, I. V. Sanotskiy, et al., 1973). More and more researchers now tend to favor the feasibility of establishing safe levels for such substances.

Figure 4 is the curve obtained in our laboratory by L. D. Katosova (1973) for the cytogenetic effects in bone marrow as a function of the chloroprene concentration. The curve has the shape of an expirant over the concentration segment studied. However, in this, as well as in other, similar investigations, only tht yield of a relative number of aberrant cells in experimental animals beyond the limits of the variations in the spontaneous level for the control, as well as for the generally accepted "norm," should be considered as the harmful effect. As is known, Gaddum (1956) considered the safe zone to be the magnitude $LD_{50} \pm 6\sigma$, so that the probability of the onset of a fatal result in this case is $10^{-9}$.

The threshold of the harmful effect of a substance is that minimum concentration of it in an environmental object, the effect of which will cause changes to occur in the organism (under concrete conditions for the admission of the substance) that exceed the limit of physiological adaptive reactions, or latent (temporarily compensated) pathology.

The commission designated to fix the maximum permissible concentrations of harmful substances in the production sphere developed a tentative definition



* – CALCULATED CURVE OF LOGARITHMIC RELATIONSHIP

$$Y = \left[\ln\left(\frac{X_n}{X} + 1\right)\right] 10,$$

WHERE $X_n$ IS THE DOSE (in mg) (REFERENCE POINT);

X IS THE MAXIMUM INEFFECTIVE DOSE IN THE EXPERIMENT, EQUAL TO 0.02 mg.

SOURCE: YANYEHAVE, N. YA., GIGIYENA I SANITARIYA, No. 7, 10., 90 (1972)

FIGURE 3. CURVE OF THE FREQUENCY OF LUNG TUMORS AS A FUNCTION OF THE DOSE OF 3, 4-BENZPYRENE INJECTED INTO THE RESPIRATORY TRACT OF MONGREL RATS.

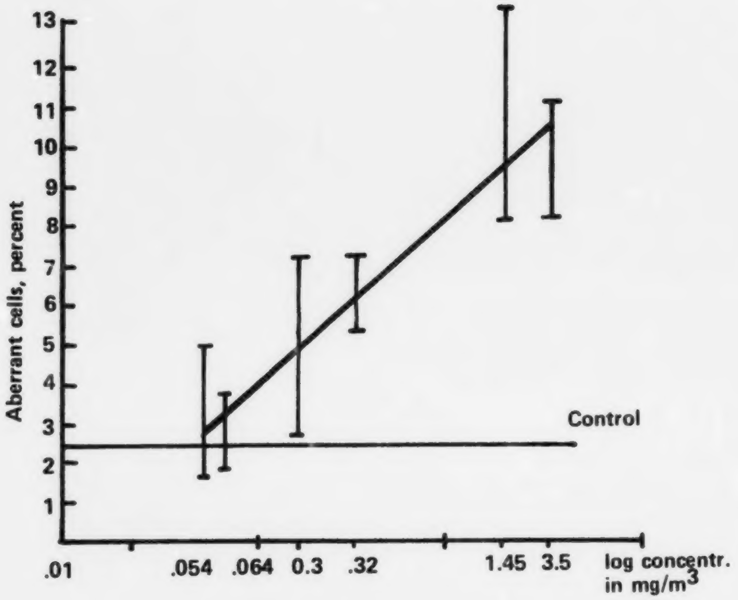


FIGURE 4. THE CYTOGENETIC EFFECT IN BONE MARROW CELLS OF ANIMALS AS A FUNCTION OF THE CHLOROPRENE CONCENTRATION

of the criteria for the *harmfulness* of changes resulting from the effect of a particular substance that was based on the departure of an index beyond the limits of annual or seasonal variations in the norm (± 2σ), as well as the stability of the changes, even though not exceeding the norm.

However, a formal approach such as this is not always satisfactory because in the first place, the norm is not known in all cases.

In our opinion, the question of criteria of harmfulness may be resolved by the delimitation of a stage of the true physiological adaptation to changed enviromental conditions and of a stage of latent, temporarily compensated pathology under the effect of minimally effective quantities of the toxin.

We have introduced into the system proposals for methodological procedures for use in resolving the problem described (I. V. Sanotskiy, N. G. Ivanov, N. M. Karamzina, V. N. Fomenko, 1971). The proposal was to use the following known biological criteria of the correspondence of environmental conditions as the basis:

- the normal condition for the population (constancy of numbers of the species);
- the capacity of the organism to adapt to the environment (the law of the unity of organism and habitat);
- preservation of the organism as a single biological system

This approach served to establish, in particular, that the use of functional and extremal loads is a delicate matter, and that their intensity must be varied widely. A comprehensive investigation is necessary to differentiate physiology and pathology in the integral organism. Thus, adaptation to stimulating toxins in terms of a respiration frequency test was accompanied by pronounced pathology in terms of morphological criteria and olfactometry (V. S. Pozdnyakov). Repairs to chromosome damage in the bone marrow as a result of the effect of prometrin were accompanied by the occurrence of chromosome rearrangements in the liver tissue (L. P. Yefimenko, A. Ye. Kulakov, 1969), and so forth.

Constancy in species numbers within known limits can be predicted under laboratory experiment conditions. The system of tests that was devised resulted in establishing thresholds of harmful effect for many toxins on the processes of reproducing future generations.

## Results

Proof of the threshold nature of the intensity of harmful effect is of enormous importance for intelligent, practical activity to protect and to improve the sanitary conditions of the environment surrounding man at work and at home, and to determine the permissible load on human populations.

If a harmful substance cannot be removed completely from production, from community and home environments, and from the biosphere in general, its content in environmental objects must be subject to toxicological limitation based on a determination of the thresholds of harmful effect. If the rate of decay and removal of the substance from the environment does not exceed the admission rate, the balance must be restored, with the thresholds of harmful effect of the substance on living systems taken into consideration. If the substance is concentrated in the environment, and this includes the food chains, its emission must be prohibited, particularly by means of closed air- and water-supply cycles.

The permissible load of a chemical factor corresponds to its maximum permissible concentration in environmental objects. This we defined (I. V. Sanotskiy, 1970) as the concentration, the effect of which on the human body periodically, or throughout a lifetime, directly or indirectly through ecological systems, as well as through possible economic losses, causes neither somatic nor psychic diseases (including latent and temporarily compensated) nor changes in the state of health exceeding the limits of adaptable physiological reactions detectable by modern research methods immediately or later in the life of present and subsequent generations.

Known is the fact that a multifactor analysis with broad formalization of the initial data is needed to make a synthetic estimate of the correspondence, or noncorrespondence, between a chemical component in the environment as a whole and hereditary, as well as acquired, properties of living systems. The most difficult stage of formalization in the case of sociological estimates is that on questionnaire methods. Nevertheless, certain publications along these lines already have been obtained.

The problem of the distribution of individual sensitivities, and of a concrete population, causes great difficulties.

Collaboration between scientists with different specialists, collaboration between scientists from different countries, is the most effective way to ensure progress in the theory and practice of protecting the biosphere against chemical pollution.

## BIBLIOGRAPHY

Akoyev, I. G.; Maksimov, G. K.; Malyshev, V M., *Luchevoye porazheniye mlekopitayushchikh i statisticheskoye modelirovaniye* [Radiation damage to mammals and statistical modeling], Moscow, 1972.

Vvedenskiy, N. Ye., *Vozbuzhdeniye, tormozheniye, narkoz* [Excitation, inhibition, narcosis], St. Petersburg, 1901.

Dubinin, N. P., *Problemy radiatsionnoy genetiki* [Problems of radiation genetics], Moscow, 1961.

Dubinin, N. P.; Tarasov, V. A., In coll: *Sovremennyye problemy radiatsionnoy genetiki* [Contemporary problems of radiation genetics], Moscow, 1969, 3.

Yefimenko, L. P.; Kulakov, A. Ye., In coll: *Mater. konfer. molodykh nauchnykh rabotnikov. Institut gigiyeny truda i profzabolevaniy* [Materials of the conference of young scientific workers. Institute of Industrial Hygiene and Occupational Diseases], 1969, Moscow, 139–140.

Isakova, G. K.; Ekshtat, B. Ya.; Kerkis, Yu. Ya., *Gigiyena i sanitariya*, 1971, No. 11, pp. 9–13.

Katosova, L. D., "An evaluation of the danger of the mutagenic effect of ethylenimine and chloroprene," Author's summary of candidate's dissertation, Moscow, 1973.

Kurlyandskaya, E. B.; Sanotskiy, I. V., *Gigiyena truda i profzabolevaniya*, 1965, 3, pp. 3–9.

Ivanov, N. G.; Kudinova, O. V.; Pozdnyakov, V. S.; Klyachkina, A. M., "Reduction of the sense of smell in animals as an index of the harmful effect of industrial toxins," In coll: *Toksikologinovykh prom. khim. v-v.* [Toxigenic industrial chemicals], No. 13, Moscow, 1973, p. 33.

Pravdin, N. S., *Rukovodstvo promyshlennoy toksikologii* [Handbook of industrial toxicology], Moscow, 1934.

Pravdin, N. S., *Metodika maloy toksikologii promyshlennykh yadov* [The methodology of minor toxicology of industrial toxins], Moscow, 1947.

Sanotskiy, I. V., "The Influence of Different Doses of Radiothorium on Reproduction," Candidate's Dissertation, Moscow, 1950.

Sanotskiy, I. V.; Avkhimenko, M. M.; Ivanov, N. G., In coll: *Toksikologiya novykh promyshlennykh khimicheskikh veshchestv* [The toxicology of new industrial chemicals], No. 9, Moscow, 1967, p. 71.

Sanotskiy, I. V., In book: *Metody opredeleniya toksichnosti i opasnosti khimicheskikh veshchestv* [Methods for determining the toxicity and danger of chemicals], Moscow, 1970.

Sanotskiy, I. V.; Ivanov, N. G.; Karamzina, N. M.; Fomenko, V. N., In coll: *Nauchnyye osnovy sovremennykh metodov gigiyenicheskogo normirovaniya khimicheskikh veshchestv v okruzhayushchey srede* [The scientific bases of modern methods for fixing hygienic norms for chemicals in the environment], Moscow, 1971, pp. 63–68.

Simonov, P. V., *Tri fazy v reaktsii organizma na vorzrastayushchiy stimul* [Three phases in the reaction of the organism to increasing stimulus], Moscow, 1962.

Strelin, G S.; *Med. radiologiya*, No. 2, 1960, p. 77.

Fomenko, V. N.; Strekalova, E. Ye.; Katosova, L. D.; Chirkova, Ye. M.; Sal'nikova, L. S.; Silant'yeva, I. V.; Yefimenko, L. G.; Kulakov, A. Ye., "Pharmacology. Chemotherapeutic agents. Toxicology." *Itogi nauki i tekhniki* [Summaries of science and technology], Vol. 5, 1973, pp. 128–145.

Hatch, T., *WHO Bulletin*, 1973, Vol. 47, No. 2, pp. 153–161.

Shabad, L. M., Sanotskiy, I. V.; Zayeva, G. N.; Bruyevich, T. S.: Katsnel'son, B. A.; Yanysheva, N. Ya.; Shugayev, B. B., *Gigiyena i sanitariya*, No. 4, 1973, pp. 78–81.

Yanysheva, N. Ya., *Gigiyena i sanitariya*, No. 7, 1972, pp. 87–91.

Yarmonenko, S. P.; Palyga, G. F., *Med. radiologiya*, No. 3, 1963, p. 66.

Bock F.C. Nat. Cancer Inst. Monogr., 1968, 28, 57–63.

Brues A.M. Science, 1958, 128, 693–699.

Dinman B. Science, 1972, 175, N. 4021.

Gaddum J. H. Brit. J. of Pharmac and Chemother, 1956, 11(2) 156–160.

Grigarzik H., Passow H. Pflugers Archiv. Bd. 267, S. 73–92 (1958).

Hatch T.E. Arch. Environ. Health, 1971, 22, 687–689.

Frohberg H., Baner A. Arzneimittel-Forsch, 1973, 23, N. 2, 230–236.

Mantel N., Heston W. E., Guriam J. J. Nat. Cancer Inst. 1961, 27, 203–215.

Mantel N. Clin Pharm. Ther. 1963, 4, 104–109.

Stokinger H. E. Arch. Environ. Health, 1972, 25, N. 3, 153–157.

Weil C .S. Toxicol and Appl. Phormocol., 1972, 21, N. 4, 454–463.

Weil C. S. Inf. Congr. Pharmacol. San. Francissa Calif. 1972. Abstrs. Invit. Presentat. (1972) 258–259.

Weisbirger J H., Weisburger E. K. Food Cosmet. Toxicol. 1968, 6, 235–242.

# Questions of Extrapolating Data in Evaluating the Mutagenic Action of Environmental Factors on Man

N. P. Bochkov *

Experiments using animals are unavoidable in evaluating the mutagenic action of environmental factors on man. No one can dispute this rule. Our success in gaining knowledge of the genetic danger to man from ionizing radiation, alkylizing compounds, and other mutagens is associated with large-scale experimental research on animals.

In the 1930's, G. Meller, N. V. Timofeyev-Ressovskiy, and other geneticists, for the first time posed the question of the danger of ionizing radiation to man's heredity based on extrapolating results obtained using animals.

Experimental genetics presently has a vast arsenal of methods for checking the mutagenic action of various factors on animals. They include cytogenetic methods applied to somatic and embryonic cells, consideration of dominant lethals, recessive lethal mutations and recessive and dominant apparent mutations. Experiments have been performed using drosophilae, mice, rats, guinea pigs, and rabbits. The most valid conclusions have been obtained from experiments using drosophilae and mice. The other animals have been used, as a rule, in cytogenetic experiments, or in experiments that took dominant lethals into consideration. Gene mutation considerations in mammals are very time consuming and expensive. Nor does consideration of the dominant lethals even begin to settle the question, because this is a collective group of genetic damages, usually vague in nature.

It is desirable, when considering the present state of the problem of extrapolating data obtained on animals to man, to discuss the following questions.

1. What are the circumstances that force us to turn to extrapolation?
2. What permits us to make extrapolations?
3. What hinders making extrapolations?
4. How should extrapolations be made?

*1. What are the circumstances that force us to turn to extrapolation?*

The enormous importance of direct data in evaluating the effect of mutagenic factors on man's heredity is well known. However, research in the mutagenic effect of environmental factors must be conducted quickly as possible. Consequently, even vast quantities of data and the most improved methods for considering gene mutations in human embryonic cells cannot accomplish the primary task, that of verifying the mutagenicity of environmental factors, since they require more time, from several years to several decades. Therefore, models must be used. These can be human cells in a culturel, or animals.

Chromosome aberrations and aneuploidy in human somatic cells can be studied by the effect of mutagenic factors on the organism, and on animals to consider chromosome and genome mutations in somatic cells. Procedures that use human cells are sufficiently well-developed for this purpose. However, similar experiments are impossible when we turn to embryonic cells. This raises the question of whether it is right to judge embryonic cells by somatic cells? Obviously not, because the formation of embryonic cells includes a number of mitotic divisions and meiosis. The nature of chromosome behavior in meiosis differs sharply from mitosis, which leads to different quantitative and qualitative behavior patterns in the mutation process in embryonic and somatic cells. This raises for the geneticist the question of the desirability of testing mutagenic factors on animals (so as to take chromosome aberrations into consideration) and extrapolating the data to man.

But it is nevertheless sometimes possible, although difficult, to obtain human embryonic cells for cytogenetic research, and thus estimate chromosome mutation frequency. It would seem that in this case we can manage without extrapolations. However, as an analysis of the data in the literature shows, there is a great difference in the frequencies of gene and chromosome mutations for the effect of the same mutagen. It is not simply a matter of the different nature of the qualitative responses, but rather that chemical mutagens can induce one type of mutation, yet have no effect on another.

---

*Institute of Medical Genetics, Academy of Medical Sciences of the USSR, Moscow

Attempting to delineate the question of gene mutations, we again are faced with the choice of one of two approaches; either induce gene mutations in a culture of human cells (the procedures have now been developed), or experiment on animals. Numerous data show the great differences in mutation frequencies *in vivo* and *in vitro*, and this is what makes experiments on animals, and extrapolation, necessary. Whether some of them are adequate, or not, will be gone into in what follows.

Thus, three circumstances force us to turn to extrapolations; the difference in the mutagenic response of somatic and embryonic cells, the difference in the frequency of induced, gene, and chromosome mutations, and the difference in the frequencies of mutations *in vivo* and *in vitro*.

### 2. *What permits us to make extrapolations?*

Many of the conditions that force us to turn to extrapolations in evaluating mutagenic factors in the environment were enumerated above. But at the same time, all are, as it were, contraindications. Can we make extrapolations anyway, or is doing so a useless task? What is it that permits us to answer the question of the correctness of making extrapolations in the affirmative?

The structure of the individual chromosomes, and of the karyotype, in higher animals conforms to a single principle. The chromosome is a strictly ordered nucleoproteinic structure, which undergoes identical changes in the process of reproduction and functioning. With the single principle of the chromosome structure as a basis, it can be assumed that there are common behavior patterns in chromosome reactions to mutagenic effects when the mutagens reach the chromosomes. Experimental data confirm these assumptions. Based on a great deal of research in the mutation process, we can speak of its universality. We can see by the example of radiation and of alkylizing agents that the similarities in the mutagenic responses of the higher organisms are greater than the differences, especially as regards the qualitative response. All substances that have been found to be mutagenic for mammals cause chromosome aberrations in human somatic cells *in vitro* and, where tested, *in vivo* as well.

### 3. *What hinders making extrapolations?*

The problem of extrapolation would not arise if there were no differences in the mutation response of the various organisms. The data obtained for any organism simply would be used to evaluate the human mutagenic danger. However, this is not so, and the more the mutation process is studied, the more the restrictions placed on extrapolation.

The first major restriction on extrapolation is associated with the difference in the metabolism of chemical substances in animals and man. Let us, by way of an example, cite thalidomide, the embryotoxic effect of which is strongly manifested in man, but not at all in rats. Many chemical compounds are toxic for one species, yet decompose rapidly in another. This most often is associated with hereditary features of metabolism.

What follows from the extrapolation restrictions considered above is that the metabolism of the mutagen in man and in animals must be known if correct extrapolations are to be made. Obviously, extrapolations will be more accurate for radiation mutagenesis than for chemical.

The second restriction placed on extrapolation is associated with the dissimilar qualitative behavior patterns in the mutation process in different animals.

Moreover, it is apparent that induced mutagenesis can be stronger or weaker in man. There is no basis for assuming that it always is more intensive in man than in animals.

What causes the quantitative differences in the induced mutation process in different animals? Induced mutagenesis, as has been established recently by different authors, is an unusually complex process, one that can be influenced by many factors. The primary damage from mutagens, apparently, is of a potential nature. The cells contain reparative systems. All of this leads to quantitative variations in the mutation process. Reliable procedures for evaluating the role of the primary and secondary processes in the differences in the mutation response remain to be developed. The importance of the phase of the cess cycle for mutation frequency, the importance of type of cells, of the species and age of the animals, has been discovered for radiation and chemical mutagenesis. The relationship between mutation frequency and mutagen dose can be different in different animals, and this fact makes extrapolation particularly difficult.

We can cite several papers to illustrate these propositions. N. I. Shapiro and Yu. Ya. Kerkis were the first to establish the fact of dissimilar radiosensitivity of the chromosomes of different species of animals.

Dominant lethals were taken into consideration in the experiments conducted by N. I. Shapiro. It was found that the numerical incidence of lethals in mice, rats, and rabbits, was dissimilar, with the rabbit showing a higher death rate of zygotes before implantation.

More comprehensive work was conducted by M. Lyon and B. Smith, who evaluated the frequencies of dominant lethals at different times after irradiation. They established the fact that it is not just a matter of the differences in the general genetic radiosensitivity, but also that the radiosensitivity will vary differently in different animals, and at different times.

The dissimilar frequency of dominant lethals in different animals and their dissimilar shift during spermatogenesis have been noted in the effect of ionizing radiation, as well as of chemical mutagens. This can be seen in the data provided by M. Lyon, B. Smith, and D. Bateman. It is important to note that the curves are dissimilar, and that the mean frequencies differ.

The differences in the frequency of induced aberrations in different organs and in different types of cells in different animals can be seen in the work done by Yu. Ya. Kerkis, et al.

The extrapolation problem is complicated by the fact that there are age differences in the response to mutagenic actions. This has been described by a number of authors for animals, and we have done so for man.

The following rule has thus far been applied to extrapolations: gene mutations increase linearly with the dose, and chromosome mutations as the square, with the genome mutations question unresolved.

These conclusions have not always been reached on the basis of strict statistical criteria. For example, it is now well-established that the dependence on the dose is not the square and determinate for radiation-induced aberrations.

The sum of the experimental data permits the conclusion to be drawn that the coefficients of the equations of mutation frequency in different species of animals (or types of cells) are different, and depend on the mutagenic factor dose. All this speaks to the enormous difficulties standing along the road of extrapolations. Things are even further complicated by the fact that the yield of muttaions per dose unit differs for different dose ranges. And as a rule, the experimenter is not working with the same doses with which man comes in contact.

### *4. How should extrapolations be made?*

The question of the mutagenicity of any factor for man is decided primarily on a qualitative basis. Extrapolations therefore, are of the same nature. What rules can we propose on the basis of the sum of contemporary knowledge?

(a) First of all, the researcher is confronted with a dilemma; should the extrapolations be made from human cells in a culture, or from animals? The evaluation of the mutation process is a very complex procedure. After all, three types of mutations (gene, chromosome, and genome) must be taken into consideration, as must at least two types of cells (somatic and embryonic), and the metabolism of the chemical mutagen. There is, therefore, no unambiguous answer to this question in all cases. The decision is different in each concrete case. Preference sometimes may be given to a culture of human cells, sometimes to animals.

(b) If mutagenicity of any factor is detected for an animal, it must be considered mutagenic for man, with the exception perhaps of those cases when the metabolism of the mutagen in animal and in man differs sharply at precisely the point of application of the mutagenic activity. Failure to observe a mutagenic effect of a compound in one, or even in several species, does not mean that it is safe for man, but only that we may hope it is. This approach seems objective to us, because of the great importance of underestimating the presence of a mutagenic effect attributable to any chemical substance (more important than overestimating). Thus, a mutagenic effect established for any animal object must be directly extrapolated to man, although this may not always be confirmed. Negative results in investigating a mutagenic effect must be extrapolated with extreme care in the sense of the absence of a mutagenic effect by a certain chemical substance on man's heredity, and only for the condition that the metabolism of the compound in the research animal and in man is exactly alike.

(c) If the mutagenic effect of any factor is established for animals, the task of quantitative extrapolation arises, because the basic approach to evaluating environmental factors must be the population approach. If numerical data are obtained for several species of animals, it is better to proceed using mean data, rather than using data on one species of animals, when evaluating the mutagenic danger to man. Of course, the task of extrapolation will be difficult if drosophilae and mice are included in this group. Preference usually should be given to mammals, particularly in the case of chemical mutagenesis. Quantitative conclusions cannot be precise, even when all extrapolation requirements are met as strictly as indicated. The estimates will be approximate in any case. Therefore one must strive to determine the order of the danger, not the precise figures. Thus, quantitative extrapolations should be made more carefully than qualitative ones.

In the case of quantitative extrapolations, it is important to evaluate the total mutagenic response, proceeding from the different sensitivity of the phases of the cell cycle, according to the weighted mean principle.

(d) The geneticist is confronted with two complex questions when evaluating the mutagenicity of a factor; what animals to use in the experiments, and which tests to select for the genetic evaluation. Obviously, the two main species of genetically well studied and convenient objects should be mentioned first, drosophilae and mice. A great many mutations can be studied in each of these species, but expanded experiments on mice to take recessive mutations into consideration are very time-consuming and expensive. The drosophila therefore is most often used. Experiments with mice as a rule are limited to dominant lethals. Experiments must use mammals when a widely distributed substance, or factor, is being checked, or when there are reasons for an assumption as to its mutagenic effect. When the mutagen does not penetrate the testicles in animals, but does so in man, experiments can be set up so that the mutagen is injected into the scrotum of the animals.

(e) It is not always necessary to attempt to make extrapolations. If experiments can be set so that human cells are used, it becomes mandatory that extrapolations not be made. There are no adequate models for evaluating chemical mutagenesis in man. Therefore, the current problem, and one that applies in all cases, is that of obtaining direct data on the influence of mutagenic factors on man's heredity. It is necessary, for example, to extrapolate the cytogenetic effects in somatic cells. Modern procedures permit doing this directly on human cells as a result of the effects of the mutagenic factors *in vivo* and *in vitro*. The large-scale development of procedures for taking gene mutations in human somatic cells into consideration will narrow the circle of questions for which it will be necessary to extrapolate data. At the same time, and in conclusion, it should be emphasized that as our knowledge of the mutation process increases, the limits of the potentials of extrapolations will become clearer, and the selection of the models used to evaluate environmental mutagenic factors will be more objective.

Extrapolation, in the final analysis, is based on the results of experimental modeling. Consequently, the better the experimental object is selected, the more accurate will be the extrapolation. Well known from experimental pathology is how attentive and careful one must be when modeling any of the human pathological processes.

# The Question of Genetic Danger from Environmental Pollutants

L. M. Filippova *

Pollution of the habitat of living organisms by biologically active agents—chemical (the wastes from different industries, erosion loss from agricultural areas, and the like), as well as physical (increase in the radiation background, thermal pollution, noise, and others)—has become a side-effect of the development of the human society.

The progressing pollution of the natural environment has a number of biological consequences, with one of the principal ones the danger of the occurrence in the majority of cases of negative hereditary changes in man, as well as in animals and plants, under the effect of chemical and physical agents in the environment. The great many chemical compounds entering the natural environment, individually and in a variety of compounds, are capable of inducing gene, genomic, and chromosome mutations, which cause an increase in the frequency of congenital deformities, defects and anomalies, many hereditary diseases, stillbirths, spantaneous abortions, and the like.

Major chromosome damage (aberrations) causes a genetic defect the scale of which is such that the damage is rarely transmitted by heredity because of the poor viability, or sterility, of the carriers, and this is precisely why aberrations have little importance from the standpoint of population.

Recessive gene mutations play a special role from the standpoint of population. They are harmful to the organism in the majority of cases. Transmitted and accumulated from generation to generation in the latent (heterozygous) state, they increase the "genetic load" of the population. The problem of the "genetic load" is a particularly serious one for a population in which the effect of natural selection has weakened, and this would apply to the human population.

Systematic study of the genetic danger of agents that pollute the environment, of products of chemical synthesis widely used in industry, agriculture, and daily life [1], has shown that many of them definitely have mutagenic activity and can damage the genetic material of living organisms. Mutagenically active compounds have been found in medicinal compounds, pesticides, dyes, and others.

Therefore, no one today can have any doubt about the fact that all chemical agents that in some way or another come in contact, in particular, with man, as well as with animals and plants, should be subject to genetic control. The results of such control should be taken into consideration in setting hygienic standards, and should further serve as the basis for drastically limiting the content of mutagenic substances in the environment. The toxicological standards accepted at the present time certainly cannot be recognized as adequate in all cases. While there are many cases when the mutagenicity of a compound correlates with toxicity, slightly toxic chemical products often have mutagenicity. And it is this mutagenicity, and not the toxicity, that should be the limiting factor when setting standards for a chemical compound such as this. Paraquinone dioxime, for example, is widely used in the tire and rubber industry and while it has no pronounced toxicity (an organoleptic index, color, is used in its standardization) it does induce lethal gene mutations and morphoses in *Drosophila* and chromosome rearrangements in a culture of peripheral blood leukocytes and in a culture of human embryonic fibroblasts [2]. Moreover, the mutagenic concentration of triethylene melamine is less than the toxic by a factor of 100 [3].

The evaluation of the genetic threat of environmental pollutants must proceed from the fact that in the majority of cases the organism will be subjected to the effects of a *large number* of mutagenic agents in *small concentrations*. There are, therefore, two problems of primary importance:

1. the problem of the simultaneous, or sequential, effect of several mutagenic agents;
2. the problem of defining the threshold of the mutation process.

1. The living organism is subjected to the action of a complex set of chemical and physical agents that penetrate the organism over different paths and from different media.

---

*Main Administration of the Hydrometeorological Service, Moscow

Recent papers have dealt with the mutagenic effect of two chemical agents, two physical (irradiation + temperature, for example) agents, or a chemical agent in combination with a physical agent (chemical mutagen + temperature, chemical mutagen + irradiation, and so on), making it possible to estimate the situation, albeit tentatively, in connection with the simultaneous presence in the environment of many mutagenic agents. The possibility of additivity, as well as of synergism of the genetic effect of two mutagenic factors, not necessarily highly active individually, has been pointed out. This increase in the level of mutagenesis also has been observed for chromosome aberrations, as well as for gene mutations [4, 5, 6]. It appears that the reasons for the intensification of the effect of combined mutagenic effect can be very different indeed. For example, the effect of one of the agents can be to change the physiological state of the cell, change the hereditary material to the labile state, damage the cell control mechanism, inhibit reparative processes, and so on, with the result a reduction in the resistance of the organism to the effects of the second agent. Nor can we exclude the possibility of the appearance of a new, more highly mutagenic substance when two compounds interact with each other, and with cell metabolites.

2. This gives rise to a question. Will the effects thus far discussed be valid for the case of extremely low concentrations of mutagenic substances in the natural environment? In other words, is there a threshold of mutagenic effect of chemical agents and radiation? It obviously is impossible to answer the question categorically. And not only because our present knowledge of the question is limited. This also is so because the genetic effect includes a complex set of damages to the hereditary material, each different from the other, and characterized by specific mechanisms and kinetics.

Shortwave radiation, and chemical substances to an even greater degree, are capable of inducing point mutations and simple, primary chromosome breaks; that is, those changes in the hereditary material the basis of which is an elementary chemical act, or a single ionization. Analysis of the kinetics of these "single-strike" mutation events attests to the linear nature of the dose-effect ratio [7, 8]. The dose-effect line can be extrapolated through the zero point, pointing to the absence of a dose threshold. Actually, the formation of a single pair of ions, or the appearance of a single molecule of a chemical mutagen in a cell, is perhaps sufficient to induce point mutation as a result of just one DNA base change. The occurrence of mutations in living organisms is probabilistic in nature. Even the tiniest doses of mutagens can induce mutation changes (which can be extremely serious, right up to being lethal, caused by a single mutation in one of millions of genes), simply among fewer persons. So, naturally enough, the probability of genetic damage will decrease with decrease in the mutagen dose. The proposition that there is no dose threshold for inducing "single-strike" mutation events is now well confirmed by direct experiments [9–13]. Those rare cases of induction of mutations attributable to the inhibition of certain enzymes needed for the production of DNA precursors are the exceptions. If such an enzyme in the norm does not limit the DNA or chromosome replication rate, the inhibitor will be effective if some minimum magnitude is exceeded [14].

But things are different in such damage to chromosome material as complex rearrangements, for which two, or more, independent chromosome breaks are necessary. Here the dose-effect curve differs from that described above. These structural changes, upon irradiation, can be caused by accumulations of ions from different tracks that occur independently. Their frequency increase is not directly proportional to the dose, but rather is more rapid, and there is a threshold dose [7]. The same thing is seen for the dose ratio in the case of the "double-strike" ("multistrike") effect of chemical agents, the result of which is the appearance of two (of several) independent breaks with subsequent regeneration and recombination of the chromosome material and the appearance of complex rearrangements of the chromosomes.

Accordingly, the kinetics of "single-strike" mutation events (the occurrence of gene mutations and of simple chromosome breaks) can be characterized by the linear dose-effect ratio and absence of a threshold. The frequency of the "multistrike" mutation processes (the occurrence of complex chromosome rearrangements) increases more rapidly than the dose, and has a threshold.

The experimental curve for the induction of mutations attributable to physical, as well as chemical, agents usually has the following, typical, shape.

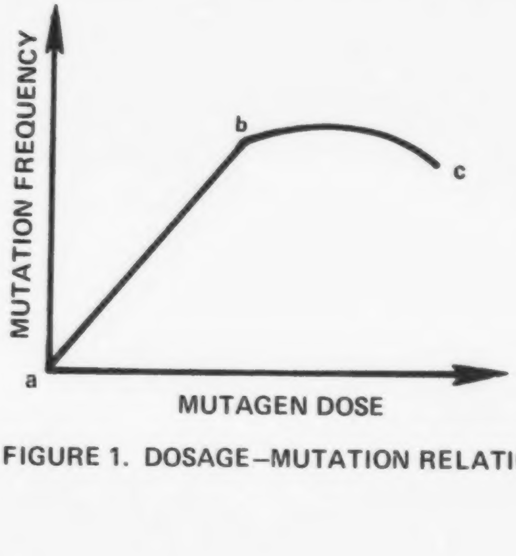


FIGURE 1. DOSAGE–MUTATION RELATION

The linear segment, ab, corresponds to the preferred induction of gene mutations under the influence of low mutagen doses. The bc segment rebects the increase in the percentage of large, complex, chromosome rearrangements and the decrease of induction of gene mutations when the mutagen doses are increased [15]. The linear segment, ab, usually is very much longer for a chemically induced mutation process than is the case for shortwave radiation. Many highly specific, active mutagens (nitroso- and diazo compounds, for example) induce gene mutations with much greater frequency than chromosome aberrations. Virtually almost all the chemical mutagens induce mainly gene mutations in the low dose range. It is obvious that the mutagenesis curve will, in practice, undergo substantial modification from compound to compound because of the high degree of specificity of the chemical mutagenesis (discussion of the nature and causes of which are pointless in this report). Nevertheless, the tendencies described will be at the bottom of the mutation process.

Some researchers assume they have every reason to talk about a threshold for toxic or teratogenic effects, as well as for mutagenic (including the induction of gene mutations). Well known is the fact that mutation frequency is largely dependent on the physiological state of the cell at the time of the action, the special features of the cell's metabolism, on a great many intra- and extra-cellular factors. Small quantities of a potentially mutagenic agent that enter the organism can be completely inactivated, and rendered harmless, by the cell's defense mechanisms, while induced pre-mutation changes can be healed by the reparative systems. These small concentrations of the mutagen thus will, for all practical purposes, be inactive. It would seem that in cases such as these there is an obvious mutagen effect threshold. But can we, at this point, say that mutagenesis has a threshold, in the strict sense of the word?

First of all, the magnitude of this "threshold" is not absolute for the same compound, and will vary greatly not only from object to object, but even from gene to gene in the same genotype. Consequently, this so-called "inactive" concentration will, to a very large extent, characterize the degree of resistance of the organism to the particular mutagenic action, and not the threshold of mutagen action. This resistance on the part of the organism is a factor that is subject to changes as it is affected by definite conditions.

Important, and convincing experimental data [16] attest to the fact that agents that are cell metabolism modifiers cause significant changes to the induced mutagenesis picture. Under real, natural conditions the actions of a great many different chemical and physical agents "subthreshold" under other conditions and concentration, and can have a definite mutagenic effect. Hydrogen peroxide, for example, will not induce mutations in many organisms, even in high concentrations, because it is decomposed by their catalase. But if a toxin that blocks catalase (potassium cyanide, or sodium azide, for example) is added the mutation frequency will increase, even when the smallest quantities of peroxide are involved.

Well known is the fact that an increase in the quantity of oxygen, carbon monoxide, hydration, visible light, temperature, and other environmental factors, can sharply increase the frequency of mutations induced by chemical mutagens and by irradiation.

There can be a sharp increase in mutagenicity as a substance metabilizes in the environment. Transformation of a substance upon entering the environment can, for example, lead to a greater capacity to be adsorbed, or to penetrate, the cell, such as is the case with inorganic mercury, which is transformed by microorganisms into methylated mercury [17].

Finally, the transformation of a "prethreshold" into a "postthreshold" dose can take place as a result of the additive, and even more so of the synergid, effect of two, and more, mutagenic agents, which we discussed above.

So, two things evidently can happen as a result of the action by small doses of mutagens.

1. The organism responds to the effect of even the smallest doses of a mutagen by straining the protective systems. This leads to inactivation, to the rapid elimination of the mutagen molecules, or to repairation of the primary damage to the genetic material. Mutation will not occur in this case because the processes involved in the mutagen action and in those opposing such action are in equilibrium. The action of the mutagen and, correspondingly, the organism's reaction are nonthreshold.

2. The strain of the protective systems reaches a maximum with increase in the genetic load. Beyond this point the action of even the smallest dose of mutagen (and not necessarily the same one) will lead to the formation of a mutation. Thus, the occurrence of a *pronounced* (pathological) change in the hereditary structures has a threshold in the majority of cases.

What follows from the above, then, is that it is of primary importance when standardizing the total genetic load of harmful substances in the environment to proceed from the nonthreshold nature of the effect of mutagenic agents, and not from the threshold nature of the occurrence of a pronounced (pathological) character. Even small quantities of a great number of chemical and physical agents in the environment are constituents in the total genetic load on the living organism. There can be little doubt that if the systems for intracell regulation and the cell's reparative systems protect the organism from the harmful effect of a small dose of a mutagen, the cell will be in no condition to withstand the simultaneous action of tens, or of hundreds, of small quantities of mutagens that would not be fatal to the organism individually. And this is understandable, because absence of effect from low doses of a mutagen is not because the low doses have no effect in the organism, it is the result of the active counteraction launched by the organism's protective systems, the capabilities of which, naturally, are quantitatively limited.

## CONCLUSION

Today, we are faced with the urgent need to expose the level of genetic damage attributable to environmental pollution. Hence, the task of genetic monitoring of the human population is an extremely important one. All cases of spontaneous abortion, anomalies of genetic etiology in the newborn, as well as in older groups, and among those exposed constantly, or periodically, to the effects of some particular agent (at work, or while undergoing treatment, for example), must be systematically taken into consideration and studied, in order to assess changes in the frequency of gene and chromosome mutations in germ and somatic cells.

We must introduce genetic methods in standard toxicological research in standardizing chemical products, and this in the immediate future. This task is well within the realm of possibility. A series of genetic studies can be made using the same animal that takes part in an acute, or a chronic, toxilogical experiment (dominant lethal mutations, chromosome aberrations in sexual, or somatic, cells, and others). Forms of genetic activity not accessible for direct analysis in mammals should be tested along parallel lines on adequate models. The determination of permissible quantities of chemical compounds in the environment should proceed from the nonthreshold kinetics of the mutation process, and, as a minimum from the summation of the genetic effect of even the smallest doses of mutagens. The permissible concentrations should be infinitely small in the case of active mutagenic agents. The only real way to control the content of these compounds is to standardize these products at the emission source, or even sooner, while they are being manufactured.

## BIBLIOGRAPHY

1. Rapoport, I. A., Filippova, L. M. *ZhVKhO im. Mendeleyeva,* 1970, 6, 681.
2. Revazova, Yu. A. "The Hygienic Importance of Research in the Genetic Activity of Industrial Compounds." Author's Dissertation, Moscow, 1967.
3. Matter, B. E., Generoso, W. M. *Abstracts.* 2d Annual Meeting of the European Environmental Mutagen Society, Czechoslovakia, 1972, p. 40.
4. Oster, I. I. In: *Advances in Radiobiology,* Oliver and Boyd (Eds.). Edinburgh, 1957, p. 475.
5. Kaufmann, B., Gay, H., Rothberg, H. *Exptl. Zool.,* 1949, III, 415.
6. Nikiforov, V. G. V kn. *Obshchaya genetika* [General Genetics]. "Nauka" Publishing House, Moscow, 1965, p. 166.
7. Gershkovich, I. *Genetika* [Genetics]. "Nauka" Publishing House, Moscow, 1968, p. 201.
8. Rapoport, I. A. V sb. *Supermutageny* [In collection: Supermutagens]. "Nauka" Publishing House, Moscow, 1966, p. 9.
9. Sparrow, A. H., Underbrink, A. G., Rossi, H. H. *Science,* 1972, 176, 916.
10. Smith, G. V sb. *Genetika raka* [In collection: Cancer Genetics]. Foreign Literature Publishing House, Moscow, 1961, p. 442.
11. Freese, E., Freese, B. *Radiat. Res., Suppl.,* 1966, 6, 97.
12. Glembotskiy, Ya. L., et al. V sb. *Radiatsionnaya genetika* [In collection: Radiation Genetics]. Academy of Sciences of the USSR Publishing House, Moscow, 1962, p. 300.
13. Marcovich, H. *These des Sci.,* 3846, 1957.
14. Freese, E. In: *The Evaluation of Chemical Mutagenicity Data in Relation to Population Risk.* Proceedings of the Workshop held at Research Triangle Park, N.C. 1973, 171.
15. Rapoport, I. A. *Khimicheskiy mutagenez* [Chemical Mutagenesis]. "Znaniye" Publishing House, Moscow, 1966, p. 24.
16. "Evaluation of Genetic Risks of Environmental Chemicals." Report of a Symposium held at Skokloster, Sweden, 1972, p. 13.
17. Miller, M. W., Berg, G. G., Eds. *Chemical Fallout.* C. C. Thomas, Springfield, Illinois, 1969. 531 pp.

# Hygienic Bases for Protecting the Environment

C. I. Sidorenko,* Ye. I. Korenevskaya,* M. A. Pinigin,** G. N. Krasovskiy ***

The problem of protecting the environment has taken on special urgency in recent years because of the scientific and technical revolution taking place in the world.

Among the many ecological, technological, and biological aspects of this problem in need of study and urgent decisions, the most important task is that of protecting human life, activities, and health (taking the interests of future generations into consideration). The hygienic aspects of the problem of protecting the environment therefore now are of paramount importance. Here at home the principle of priority in protecting health, and the preservation of favorable hygienic living conditions for the population, was clearly expressed in a number of resolutions by the Central Committee of the Communist Party of the Soviet Union, the Supreme Soviet, and the Council of Ministers of the USSR on improving protection of the environment and on the rational use of natural resources.

Wide-ranging governmental measures aimed at preventing pollution and sanitation of the environment in the interests of the population are based on hygienic norms that set a maximum permissible content of harmful substances in water, atmospheric air, and food products. A hygienic norm, or the maximum permissible concentration (MPC), is that concentration of a chemical substance the observance of which results in no unfavorable effect on the health, well-being, or working capacity of the population, and of future generations, as detectable by modern research methods, while at the same time causing no deterioration in the population's hygienic living conditions.

The development of the principles, and of the methodology, for establishing the MPC is a major achievement of Soviet hygiene, the more so because Soviet scientists (N. V. Lazarev, A. N. Sysin, S. N. Cherkinskiy, and V. A. Ryazanov) made the basic contribution to the development of this scientific trend. The State Sanitation Code of the USSR, which is designed to protect man's environment, now includes over 420 norms for atmospheric pollution that involve the isolated and combined effects of these substances.

Setting a hygienic norm for chemical factors in the environment is based on two main principles. One is the concept of the presence of biological thresholds in the effect of these factors. The other is the principle of a limiting harmfulness criterion. Thus, when the MPC of chemical substances in the waters of reservoirs is scientifically substantiated, complex experimental research is conducted, resulting in the establishment of (a) the threshold concentration in terms of influence on the organoleptic characteristics of the water (odor, taste, color, turbidity), (b) the threshold concentration in terms of the influence on the overall sanitary regime in the reservoir (for which only initial changes in the biochemical processes of the mineralization of organic substances in reservoirs are detected), (c) the ineffective (subthreshold) concentration, which has no unfavorable toxic effect on the human body.

The concrete value of the MPC is established in terms of the most sensitive index (the harmfulness criterion), characterized by the least magnitude.

The USSR has established two norms for atmospheric pollution: the maximum single, and the mean daily, MPC. The maximum single MPC has been established to prevent reflex reactions in man during short-term effects of atmospheric pollution, while the mean daily MPC has been established to prevent its resorptive effect.

Experimental research is the basis for setting hygienic norms in the field of general and municipal hygiene. The importance of experiments stems:

first, from the interests of preventive sanitary inspection, a procedure that can implement measures against chemical pollution of the environment during the planning stage for new enterprises and plants, and thereby prevent the unfavorable effect of such pollution on the population;

second, from the possibility under experimental conditions of revealing relatively accurately the relationship between each biological effect and the level of concentration of the chemical substances, something that is difficult to do under natural conditions because of the additional effects of all the different factors present in man's habitat.

---

*Professor
**Docent
***Senior Scientist

Experimental research in establishing safe levels of chemical pollution for man is proceeding in two versions:

(a) an experiment, using volunteers, to reveal threshold and subthreshold concentrations by reflex reactions during brief periods of inhalation of a substance in terms of its influence on the organoleptic characteristics of water that is completely safe for the subjects;

(b) a lengthy, continuous experiment on animals to establish threshold and subthreshold concentrations in terms of a toxicological criterion of harmfulness of the substance.

The first stage in the development of hygienic normalization in the USSR (the 50s through the 70s) has been characterized by the accumulation of a great deal of factual data to justify the MPC for those chemical factors which are most often the cause of environmental pollution. Theoretical generalizations of the accumulated data have begun to be made in recent years because of the need to carry out the important task of increasing the reliability of the MPC that have been established by perfecting the methodological and theoretical bases for hygienic normalization.

The problem of the combined and complex effect of substances is deserving of special attention. We know that mixtures (combinations) of substances can cause different effects, synergism, antagonism, an independent effect, depending on the nature of their toxic effect and on content level. The experimental research that has been conducted to estimate the nature of the combined effect of substances polluting the air, as well as the reservoirs, has shown that an additive effect occurred in many cases. However, this question still is far from resolved because of the extreme complexity of the composition of pollutants, and the relatively low level of their content in the environment.

This means that we must make further improvement in the methods used to study the combined effect when concentrations of substances are at low levels, as well as to study the principles for estimating the nature of the summed biological effect of complex mixtures. Experimental research is now in progress along just such lines in institutes in the USSR.

Even more complicated is estimating the danger of the complex effect of a substance entering the body simultaneously over different pathways (from different media). The current practice of establishing hygienic norms for the same chemical substances separately for atmospheric air, the air in production areas, for water in reservoirs, and for food products fails to consider the possibility of their entering the body simultaneously over different pathways. This gives rise to the need to develop methodological approaches to complex (unified) hygienic normalization for chemical compounds in the environment.

In view of the difference in the principles and criteria for establishing maximum permissible concentrations of harmful substances for different media, what we are talking about at this stage is the development of approaches to unified hygienic normalization for substances, the permissible content of which can be limited by their resorptive, that is, their toxic, effect.

The integration of hygienic normalization, or unified hygienic normalization, in different environmental objects, while an urgent requirement of the times, is an extremely complex problem, the solution of which certainly will not be a unique one in all cases.

We know that a different content often finds its way into the very concept of "unified hygienic normalization."

It is not our purpose to provide a precise and universal formulation of this concept. We will simply point out that at this time we are not including in this concept the task of establishing a unified MPC magnitude for all media. The problem of integration of hygienic normalization is associated specifically with answers to the following questions:

1. The development of principles and methods for estimating the danger and safety of harmful substances upon their complex admission to the body, as well as unified approaches to estimating the simultaneous pollution of different environmental objects.

2. The development of unified criteria for the harmfulness and harmlessness of a chemical factor for the body, and scientific justification for unified principles and conditions for conducting toxicological experiments (length of experiments, selection of doses of substances, research methods, interpretation of results obtained, selection of safety factors, and so forth).

3. The carrying out of organizational measures providing for unified planning of scientific research in the study of toxic substances as applicable to all media and the simultaneous review of data obtained by representatives of the different branches of science.

The first step could, for example, be the establishment of unified norms protecting the interests of all branches of the national economy in the rational use of reservoirs (ensuring sanitary conditions of water use by the population, the interests of fish protection, the preservation of hydrobionts, the use of reservoirs for watering agricultural crops, and so forth). Accumulated practical experience in the use of sanitary and fishery MPC, with the orientation being toward a more rigid norm from these two standards, has shown the desirability and soundness of this approach. Unfortunately, future realization of this task has been made difficult by the lack of methodological bases for the normalization of MPC magnitudes in hydrobiology and agriculture.

There probably are different ways to estimate the danger from chemical substances upon their complex admission to the body. However, the establishment of the biological equivalency (the isoeffectiveness) of doses, and of concentrations, of substances entering the body over different pathways is, as we see it, of decisive importance in developing approaches to unified hygienic normalization.

Various estimates have it that under contemporary conditions, human health is threatened by several thousand chemical substances, very few of which have been studied, relatively speaking. At the same time, knowledge of the degree of danger to be anticipated from each environmental pollutant is extremely important.

Despite the success achieved in the field of hygienic normalization for environmental pollution, the rates at which norms are validated lag sharply behind practical requirements. This is what increases the urgency for scientific development of methods for the accelerated establishment of MPC for chemical pollutions.

Accelerated normalization for harmful substances can be accomplished by developing calculating and rapid methods. Formulas for predicating MPC in atmospheric air, and in the waters in reservoirs, in terms of the physico-chemical characteristics of the substances, of the parameters of their toxicity, and of the norms established for other media, that are based on correlation and regression analyses, have been proposed.

Another approach to accelerated validation of hygienic norms is based on a study of the "concentration-time" and "dose-effect" relationships in a short experiment with subsequent extrapolation of the data obtained to an extended period of time.

Research to validate the methodological conditions for the practical use of such an approach is presently in progress along these lines at the A. N. Sysin Institute of General and Municipal Hygiene of the Academy of Medical Sciences of the USSR.

Estimating the danger from chemical compounds, with consideration given to their intermittent effect, is an extremely urgent problem in hygienic normalization. We know that the concentration of harmful substances in the environment is constantly undergoing temporal change, and as experimental research shows, the intermittent effect can, in some cases diminish, and in others strengthen toxic effects as compared to the effect of permanent concentrations.

It is extremely important to the estimate of the intermittent effect to establish the general behavior patterns in the development of the toxic effect for both the permanent and interrupted effects of chemical compounds, because it is a practical impossibility, given experimental conditions, to reproduce the picture of the variability of the concentrations of the substances polluting the environment. First among these general questions that should be included are the quantitative dependence of the biological effect on the concentration of the substance and the duration of its effect, comparative estimates of the permanent and interrupted effect of harmful substances on the body, quantitative estimates of the capability of the substances and the cumulation of the effect, and so-forth.

Wide-ranging special research in validating the principles and methods used to study the long-term consequences (blastomogenic, genetic, gonadotropic, teratogenic, allergenic) of the effect of chemical pollution is needed to resolve the question of the feasibility of using these data for hygienic normalization. Questions of the threshold nature of the types of biological effects of harmful chemical substances indicated still have not been developed sufficiently in the theory of hygienic normalization. The question of the feasibility of normalizing carcinogens in the environment now has been resolved favorably, in principle (on the basis of the MPC we can take a dose that will "cause" an effect beyond the limits of human life), but the feasibility of normalizing mutagens and allergens still is under discussion. Hygienic science must develop this most important field of research more effectively.

One of the key questions of hygienic normalization involves the problem of extrapolating toxicological data from animals to man. A toxicological experiment can be called biological modeling, because there can be no identity between the bodies of animals and man, and it is this that makes it necessary to resolve the question of the conditions under which experimental data can be put into hygienic practice. The extrapolation methods used in hygienic research (direct transfer of data, the use of safety factors, and so forth) are quite empirical in nature.

It has been established that the method that makes use of direct transfer of toxicological data from animals, from white rats, for example, to man must be used with care in order to validate hygienic norms. The desirability of using "universal" safety factors also raises doubts, because the value of the coefficient of extrapolation must be found in terms of the species of laboratory animals, and of the specific properties of the toxin being studied. The method used to convert doses per unit of body surface of animals and man in order to grade species differences with the toxins also is in need of correlation.

The validation of the conditions under which experimental data can be extrapolated from animals to man is attended by great difficulties, but we believe that study of questions concerned with the comparative sensitivity of man and animals to toxins, and the recognition of general behavior patterns in species differences between mammals toward the effect of chemical pollution, has to be considered a promising approach to resolving this complex problem.

Thus, the development of methodological and theoretical research is the most important prerequisite for increasing the scientific validation and reliability of hygienic norms. It is no accident that in accordance with the program for scientific collaboration between the USSR and the USA on the problem of "Environmental Hygiene," the combined efforts of the scientists of both countries are concentrated on finding answers to precisely these current questions of hygienic normalization. The first results of the joint research show that we now have great potential possibilities for creating an even more perfect scientific methodology for perfecting legislation in the field of protecting the environment, which is the basis for initiating measures that will improve hygienic conditions.

# Hygienic Normalization Under Scientific and Technical Progress Conditions

A. P. Shitskova, M. I. Gusev,* S. M. Pavlenko, I. L. Karagodina

Scientific and technical progress is sharply changing the nature of the effect of society on nature, and is bringing about changes in the biosphere. The primary factors affecting the biosphere are urbanization and industrial, scientific, and technical progress. They cause change to take place in the environment, in atmospheric air, water, soil, nutrition, and in other things, and the appearance of new factors, such as vibration, ultrasonics, static electricity, electromagnetic oscillations, and others.

Control of the biosphere under these conditions is one of the most important tasks confronting mankind.

The classics of Marxism-Leninism considered this problem to be a social one, right from the very beginning. Control of the biosphere is a conscious activity on the part of society, aimed at regulating life processes, including protecting the health and the biological perfection of existing populations, as well as of future generations.

The implementation of measures designed to improve the environment requires the development of quantitative indices for the condition of different environmental factors and criteria of harmlessness that ensure optimum conditions for man's (society's) vital activities under predetermined social conditions.

Hygienic normalization during a period of rapid technical development is universal for all environmental objects, and is brought about by practical demands. Scientifically developed quantitative indices (maximum permissible concentration, maximum permissible value, maximum permissible level, and others) were until recently concerned mainly with individual substances and environmental factors.

Research in the course of hygienic normalization for harmful substances includes comprehensive experimental research methods on animals, observations of people, and natural investigations of the state of the environment.

The whole arsenal of modern physiological, biochemical, pathophysical, physicochemical, clinical, and other research methods is used to study the biological effect of small concentrations of toxic substances in order to provide hygienic normalization for these latter. Hygienic normalization, as a rule, is aimed at establishing threshold and subthreshold concentrations of the substances or environmental factors studied.

The estimate of the interaction between toxic substances and the organism depends on the dose-time-effect ratio. This means that all environmental factors can become pathogenic if their duration and intensity exceed the threshold level of the effect.

The principles of hygienic normalization have been perfected with the development of hygienic science and new tasks, such as recognition of the consequences of multifactor environmental effects on the organism, have arisen before the researchers in the modern stage. As a result, in addition to establishing permissible limits for individual environmental factors, questions concerned with the study of the combined effect of toxic substances on the organism have become currently important. We know that the simultaneous effect of toxic substance on the organism can be the same as the effect of each individually (independent action), can be weaker (antagonism), or stronger (synergism). Simple summation (additive action) can take place in the latter case, when the sum of the concentrations of two or more toxic substances with a definite effect is equal to unity.

Based on present data, and confirmed by experimental research in the field of municipal hygiene for atmospheric air and water in reservoirs, there usually is an additive effect, that is, the effect of summation, in the case of the combined effect of toxic substances.

It is the opinion of I. V. Sanotskiy that potentiation is extremely rare in the field of industrial hygiene when there is a lingering combination of the effect at the level of the threshold of lingering effect.

Thus, the accumulated data from the experimental study of the nature of the combined effect of harmful substances in low concentrations entering the organism by whatever pathway have made it possible to formulate principles for the normalization of combinations of substances included in the corresponding legislative resolutions.

---

*Institute of Hygiene, Moscow

Hygiene has recently had to study the combined effect of physical and chemical factors under production conditions and in populated areas.

The best studied of the physical factors include noise, vibration, ultrasonics, and temperature variations. Their combinations with chemical substances can be quite varied. The combined effect of physical and chemical factors differs from their isolated effect on the organism. Thus, clinical and experimental research has established that local and general vibration intensifies the toxic effect of heavy metals (N. A. Gordonova, 1967; L. Ya. Tartakovskaya, 1970).

Experiments on animals have shown that general vibration combined with manganese, mercury, or lead has a more pronounced effect on some indices of the functional state of the central nervous system, metabolism, and others. New, distinctive, phenomena not noted when the effect is isolated sometimes will be observed. However, changes pointing to the attenuation of the mutual effects of different agents can be detected, along with potentiation. Thus, vibration definitely moderated the anemia-causing action of mercury or lead, showing a direct effect on the hematopoietic processes in the bone marrow (L. Ya. Tartakovskaya, Ye. N. Fresh, A. G. Mal'kova, 1972).

Manganese exhibits less of an effect on the delicate structure of the peripheral nerve-muscle endings, and the regeneration of damaged elements begins much more quickly as compared to the isolated action of manganese. The data provided by Ye. M. Neizvestnaya, L. G. Babushkina, and A. G. Gol'del'man (1972) show attenuation of the mutual effects of vibration and $MnC_2$, as well as of silicon in lung diseases.

The combined effect of ultrasonics and ethanol intensifies the biological effect obtained from the isolated effect of these factors. Changes in the central nervous systems of animals under the repeated effect of this complex under study were of the same nature as they were under the effect of ultrasonics alone, but in a more pronounced form (A. D. Shipachyova, 1969). This leads to the assumption that the physical factor, that is, the ultrasonics, is the lead component in this complex. Thus, reactions of the organism to the combined effect are ambiguous. The reactivity of organs, and of functional systems, at different levels can be distinguished by their special features, and the integral biological effect is not the result of simple arithmetic addition, but reflects the specific differences between the reactive structure of the organism.

Certain common "critical" functional systems which, in the final analysis, determine the relationship between the processes of pathogenesis in the course of the response reaction, also can be distinguished.

Intensive scientific and technical progress has greatly complicated the questions confronting hygiene in recent years. The need arose to develop new methodological approaches to making a complex estimate of the effect of environmental factors. Today, research is in progress along these lines to arrive at complex hygienic normalization of pesticides in environmental objects (L. I. Metdved', Ye. I. Spynu, et al., 1969). Some calculations of the maximum harmless dose for the simultaneous entry of substances from different media have been proposed (A. I. Korbakova, N. I. Shumskaya, et al., 1971).

One of the possible methodological variants in developing approaches to complex hygienic normalization can be the use of graphic calculation methods (N. I. Sidorenko, N. A. Pinigin, 1971, 1972).

Attention also is being given to the experimental study of the special features of the complex effect of toxic substances in the environment. The complex effect of the substances cannot be equated to the combined effect without the corresponding experimental verification. It can be assumed that the complex effect will have its own special features. This is so primarily because the same chemical substances will act on the different receptor fields in a manner that will depend on the pathway over which they entered the body (R. V. Anichkov, 1955; A. G. Bukhtiyarov, 1955), and will differ in the nature and rate of adsorption, conversion, and excretion from the body (Ya. A. Teysinger, 1959; I. L. Gadaskina, et al., 1971).

Special experimental studies were made at the F. F. Erisman Institute to study the general behavior patterns of this effect by the most widely used chemical substances found in organic synthesis production, such as ethanol, cyclohexanone, benzene, methanol, formaldehyde, carbon tetrachloride, and other nonelectrolytic toxins (S. M. Pavlenko, V. A. Guseva).

The scheme developed for studying the complex effect of substances in a continuing experiment provided for revealing the effect of the toxins at the level of threshold and inactive doses and concentrations. Maximum attention was given to the selection of common, pathogenetically meaningful, and the most sensitive, specific tests and functional loads in order to seek out the special features of the effect of the same substances when body entry was over peroral, inhalational, and complex pathways.

It has been shown by the example of the study made of the behavior patterns involved in intoxication by nonelectrolytic substances that at the basis of the origin of their chronic intoxication are mechanisms that affect the central nervous system, as well as some of the metabolic processes, regardless of the entry pathway, although the formation of chronic intoxication did proceed differently. All response reactions occurred earlier for most of the substances when entry was via a complex pathway than when the pathway was inhalational or peroral.

However, the toxicity of the nonelectrolytes studied changed unidirectionally upon complex introduction in the continuous experiment, despite the fact that benzene, alcohols, cyclohexanone, and other of the substances studied are opposite in type of effect from industrial narcotic-toxins, which belong to different groups in the system of nonelectrolytes.

We, in our studies, paid special attention to discovering the breakdown in the body's adaptive mechanisms (adaptations) to the effect of the toxic substances studied when they entered via different pathways, using a set of functional tests. A comparative estimate of a series of integral and specific functional loads (static efficiency, a cold test, an alcohol test, injection of a decisive dose of the substance itself, anodization of the brain, and others) was made for this purpose.

The experiments showed that in the presence of even negligible symptoms of intoxication, all the functional loads used yield a good effect, revealing any change in the functional activity of the organism. So far as the stage of apparent normalization is concerned, only specific pathogenetically meaningful functional loads reveal the latent effect of the toxin.

The specific functional test for substances with a narcotic effect is anodization of the brain.

The studies made showed that the formation of chronic intoxication proceeded differently when industrial nonelectrolytic toxins at the threshold level entered over different pathways.

Basic changes in the indices for the functional state of the organs and systems were noted only during the first three to four months of exposure when entry of substances was complex, as opposed to separately (inhalation or peroral).

The state of nonspecifically increased resistivity appears in the first weeks of the experiment. The stage of compensation for the pathological process most usually takes three to four months to develop. Data also were obtained which indicate the possibility of the development in this case of latent phases of intoxication from the toxins studied in shorter periods of time as compared to the separate effects of these same substances. This fact is of scientific interest, because it makes it possible to correctly estimate the danger of chronic intoxication, with the complex effect of environmental factors taken into consideration.

Study of the intoxication mechanisms makes it possible to approach the development of unified methods for estimating the biological effect of toxins (nonelectrolytes) when they enter over different pathways (peroral, inhalational, and complex).

Generalization of the experimental data helps reveal the general behavior pattern in the reaction of the organism to the simultaneous entrance of harmful organic substances (nonelectrolytes) over different pathways, and usually is expressed in the effect of the summing of their effect.

Along with the study of the complex and combined effects of substances on the organism, attention recently has been concentrated on the development of methodological questions with a view to ascertaining the behavior patterns that could lead to integrating the MPC for toxic substances simultaneously in food products, water in reservoirs, atmospheric air, and other environmental objects. Calculation methods based on ascertaining the correlation relationship between different parameters of toxicity and chemical constants were developed in this connection.

The development of the multifactor relationship between the parameters of toxicity of acute and chronic effects and the physicochemical indices for substances normed in food products, water in reservoirs, atmospheric air, and in the area of working spaces, is the basis of mathematical modeling research.

The use of parametric and nonparametric analysis methods has made it possible to expose the correlation relationship between the above-indicated indices, based on which calculation formulas have been obtained for use in finding the tentative value of the permissible content of toxins in different media.

On the other hand, the development of accelerated methods for establishing the MPC for substances in various environmental objects, based on a study of the special features of the biological effect of toxins and the exposure of common pathogenetically meaningful indices of toxicity for different entry pathways, was begun in a combined experiment. This approach permits the development of unified methods for estimating the biological effect of chemical substances and makes it possible to conduct studies simultaneously, and in parallel, that deal with hygienic normalization of toxic substances in different media.

Our data show that predicating the harmless level of chemical compounds in different environmental objects while combining the use of mathematical modeling and detailed study of the behavior patterns of the biological effect of toxic substances is a promising methodological direction to take in the integral approach to hygienic normalization.

Scientific and technical progress, created by man, cannot be considered as something that is alien to the biosphere. It simply is a qualitatively new stage in its development, one that can be regulated from the point of view of the interrelations between man and nature. Therefore, we ought not consider the presence of changes occurring in the environment as a fatal inevitability accompanied only by negative effects on the organism. It would be more correct to consider environmental pollution (of atmospheric air, water, soil, and others) as a phenomenon occurring as a result of what still are imperfect technological processes. All the conditions needed to perfect these processes are available in the socialist society. Mankind today is at the stage of a gradual, intelligent, transformation of the biosphere into a zoosphere. Therefore, the activity of the hygienists during the period of the scientific and technical revolution should be directed at studying the effect of the newly-forming environmental factors on the organism on a new scientific methodological base. This permits developing multiprofile complex indices in accordance with which it will be possible to correlate the permissible parameters for forecasting the environment, working conditions, living conditions, and the health of the population; to protect man's health heredity against the defective action of mutations; to preserve existing genetic information; and, by social reorganization, to promote the creation of a healthy, modern, society.

In his report to the 24th Congress of the CPSU, L. I. Brezhnev pointed out that the consolidation of the achievements of scientific and technical progress and of socialist management of the economy in our country permits the establishment of rational forms between nature and society.

# Maximum Permissible Human Stress*

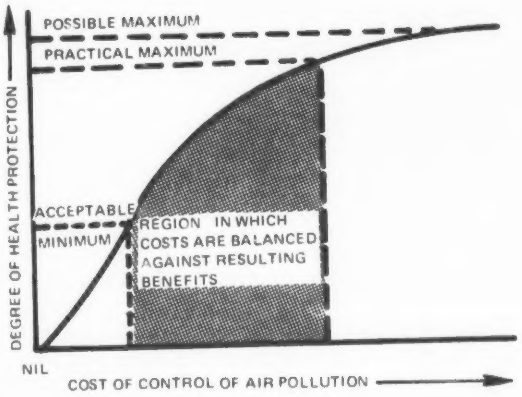
Vaun A. Newill **

## I. INTRODUCTION

Formulation of sound environmental policies depends upon the availability of adequate information upon which to base regulatory decisions. Current environmental health research must provide data for adequate assessment of health risks that arise from environmental pollutant exposure. Pollution abatement to reduce or negate these health risks requires financial expenditures and possibly even lifestyle changes. Figure 1 is a simplified schematic diagram that presents several of the principal factors in the decision-making process concerning air pollution control. It indicates that the degree of health protection attained is a function of pollution control costs. The minimum acceptable level of health protection is, at the very least, that level necessary to protect from death; and the immediate regulatory action program adopted for air pollution control should certainly protect from illness as well.[1]

The degree of health protection to be selected above the minimum acceptable level from Figure 1 is a matter for political decision. The appropriate authorities must decide on the level of health protection desirable for their society. Increments of health protection above the minimum acceptable level are generally purchased at ever increasing increments in control costs. The zone in which increased health protection (benefit) is obtained at increasing control costs (the cross hatched area in Figure 1) can also be thought of as the region of social decision making. The level of health protection desired must, of course, take into account the existing air pollution effects, but other considerations are also important, including general social, cultural and economic factors, as well as the magnitude of other health problems. In deciding where to spend limited money resources, the total set of risks that man must bear need to be considered and the greatest efforts should be expended to reduce the largest risks, be they environmental or other societal problems.

Each regulatory action is taken for society's benefit. However, each action also has a societal disbenefit because it consumes a certain quantity of the society's resources for its accomplishment. Obviously there must be some give and take between the societal benefits and disbenefits. Decisions involving this process are difficult because the group to benefit may be different from the group that receives the disbenefits and thus, it is difficult to make this an equitable process.



SOURCE BASED ON WLD. HLTH. ORG. TECHN. REP. SER., 1972, No. 506

FIGURE 1. DEGREE OF HEALTH PROTECTION VS. COST OF AIR POLLUTION CONTROL.

Another important concept follows from further consideration of Figure 1. Infinite health protection is unattainable, thus society must either decide what level of health risk is acceptable, a decision that will be reflected in the societal financial investment in environmental pollution control; or the decision will be made indirectly since the societal financial investment in environmental pollution control will dictate the level of health protection to be achieved.

## II. ENVIRONMENTAL PROBLEM ASSESSMENT

Environmental problems are identified in two general ways. One is by focusing on the environmental factors (agents) to which the population is exposed.

---

*Much of the content of this paper is based on internal EPA reports. Only those from which many passages were taken are cited. References other than EPA reports are cited.

**Special Assistant to the Administrator, U.S. Environmental Protection Agency

Another is by focusing on disease determinants and mechanisms.

By such considerations many problem areas are identified. To decide which problem areas are the most important some criteria must be applied. Factors considered in placing the problems into a priority order for action include such considerations as: (1) characteristics of the pollutant, e.g., ubiquity, expected levels of exposure to the population, innate toxicity, bioaccumulation and persistence; and (2) the pollutant's potential for playing an important role in public health problems, e.g., affecting the frequency, severity or trend pattern of a specific disease. Pollutants suspected as being related to diseases that are common, severe and increasing in prevalance, should gain the highest priority for regulatory action.[2]

Now let's consider some issues in the detailed assessment of specific problems.

Integrated information from several disciplines is useful in the health assessment of these problems. The major disciplines involved are epidemiology, clinical research and the whole set of experimental studies that will be included under the term toxicology, being fully aware that many other disciplines are involved in these studies.

Each of these discipline approaches is useful. A well-designed program ,integrating the unique capabilities and the crosswalks between these disciplines can provide a most satisfactory basis for regulatory action. Epidemiology can provide studies of population (communities or disease groups, current or long-term) exposure in real life settings. The advantages of epidemiology are: the exposure is natural, there is no need for extrapolation of data to the human, the most vulnerable groups in the population can be studied, and both current and long-term low-level exposures can be evaluated. The major problems relate to quantifying the exposure, to dealing with the many covariates, to obtaining dose-response data and to deciding association vs. causation.

Clinical research studies can be used to gather human data on either normal or diseased persons regarding absorption, metabolism and excretion of pollutants and can be used for indepth studies of humans accidently exposed to high levels of pollutants. These accidently-exposed humans provide an opportunity to develop new study parameters and response indicators. The advantages of clinical studies are: the pollutant exposure is controlled so that improved dose measurements are obtained; since each person can usually be used as his own control, covariates are well controlled; vulnerable (highly susceptible) subjects can be included in the studies; cause-effect relationships are more easily ascertained; there is no need for extrapolation to humans; and thus the derived information gives maximum input into standards. The problems are that the exposure is artificial, there can be no long-term exposures and thus only acute effects are determined, and there can be real hazard to the exposed persons.

Toxicologic studies can use many response systems, such as whole animal, isolated organs, cells or biochemical systems. The advantages of toxicologic studies are: maximum dose-response data can be obtained, though this information is incomplete at the low end of the curve; data acquisition is rapid; cause-effect relationships are more sure; and mechanism of response studies, such as kinetics of pollutant absorption, distribution, metabolism and excretion, can be performed. While known quantitative exposure requirements are most easily satisfied under controlled or experimental exposure settings, appropriate human disease models in animals have not been available and thus have not been evaluated in this manner. Furthermore, laboratory experimental studies cannot provide complete assurance of dose effects because community exposures cannot be duplicated in the laboratory. In addition, the difficulties of extrapolating the data to the human remain, particularly estimating the threshold of human response.

## Pollutants, Response and Human Exposure

"The effects of pollutants on human health depend on the physical and chemical properties of the pollutant, on the duration, concentration and route of exposure and on the human uptake and metabolism of the pollutant. Man's biological response is likewise a function of occupational, psychosocial and climatologic factors and is tempered by the phenomena of tolerance and adaptation. These exposure factors underlie attempts to understand the impact of pollutant exposure on human health.

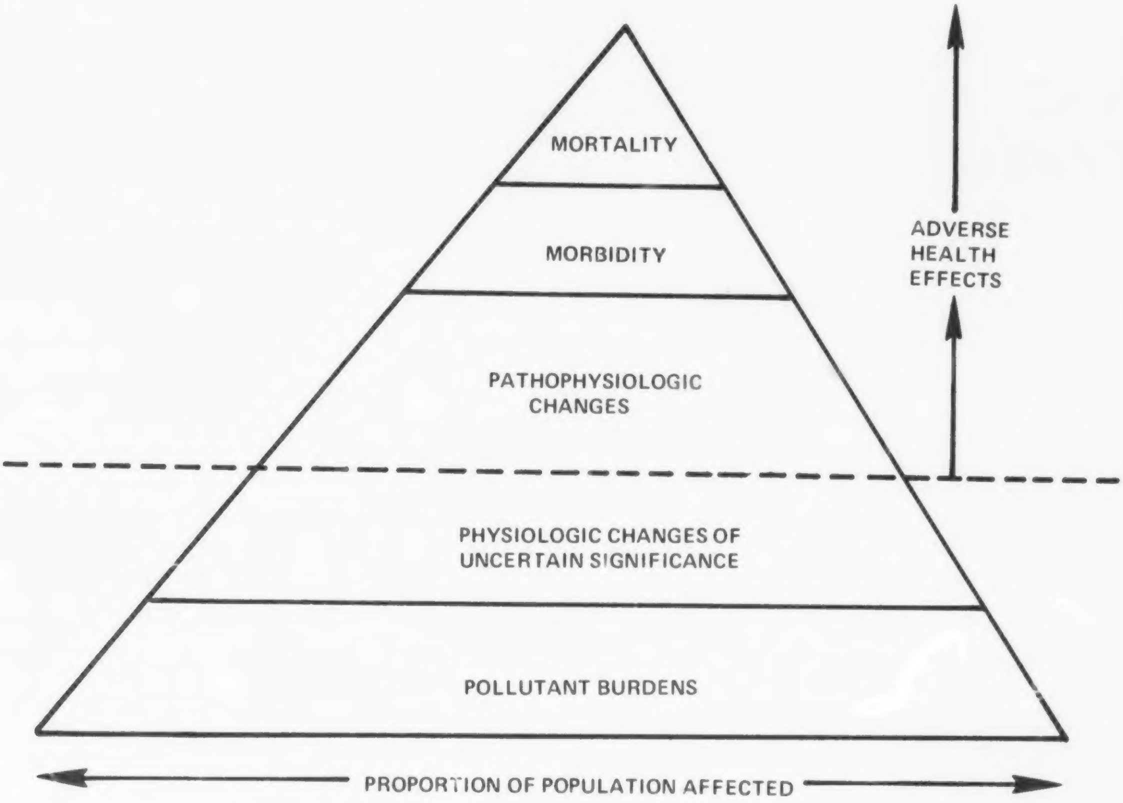
"The physical and chemical properties of pollutants determine their potential as a health hazard. These properties—including size, density, viscosity, shape, electrical charge, volatility, solubility and chemical reactivity—all affect the absorption, retention and toxicity of the pollutants. Many pollutants do not retain their original identities after entering the environment. Thermal, chemical and photochemical reactions occur when pollutants move through the environment from source to receptor. These factors affect the final physical and chemical state at the point of human exposure and help determine the toxic potential of the pollutants."[3]

## The Response Spectrum[3]

Environmental pollutants can affect the health of individuals or communities over a broad range of biological responses. One can conveniently think of five biological response stages of increasing severity as illustrated in Figure 2: (1) a tissue pollutant burden not associated with other biological changes, (2) physiologic or metabolic changes of uncertain significance, (3) physiologic or metabolic changes that are clear-cut disease sentinels, (4) morbidity or disease, and (5) mortality. Boundaries between categories may occasionally overlap. Furthermore, each category shows a range of responses rather than a simple all-or-none phenomenon.

At any point in time more severe effects, such as death or chronic disease, will be manifest in relatively small proportions of the population. In very few cases can death or disease be attributed directly and solely to pollutant exposure. Death and disease are end products of repeated cumulative insults (cumulative risks) from sources such as diet, cigarette smoking, physical inactivity, infectious challenges, and accidental injury. In general, the role of environmental contaminants in the mortality or morbidity experience of a community is difficult to quantify because so many other determinants of death and disease cannot be adequately measured.

The lower levels of the response spectrum shown in Figure 2 are subclinical manifestations of pollutant exposures. Larger portions of the population are affected to these levels of response. Pathophysiologic responses such as impaired mucociliary clearance and bronchoconstriction, and physiologic changes of uncertain significance, such as neurobehavoral responses, are more adaptable to experimental studies on animals or humans than is the case with acute or chronic disease and can be more readily associated with specific pollutant exposures. Pollutant burdens are tissue residues resulting from pollutant exposure. Pollutant burdens are highly specific effects of exposure, can be readily quantified in population studies and may be used as indicators of environmental quality. If the bridge between the lower and higher levels of the response spectrum can be established, the disease risk associated with pollutant burdens or physiologic changes can be shown, and ultimately the role of pollutant exposure in the total community morbidity and mortality experience can be defined.



SOURCE: PROCEEDINGS OF THE CONFERENCE ON HEALTH EFFECTS OF AIR POLLUTANTS, OCT. 3-5, 1973; COMMITTEE ON PUBLIC WORKS, UNITED STATES SENATE 93-15; p. 646.

FIGURE 2. SPECTRUM OF BIOLOGICAL RESPONSE TO POLLUTANT EXPOSURE

Some groups within the population may be especially susceptible to environmental factors. Notably these include the very young, the very old, and those affected by a disease. Susceptibility, however, may be temporary as well as permanent. Inherited abnormalities such as alpha antitrypsin deficiency and abnormal hemoglobins are examples of permanently altered sensitivity. Temporary increased sensitivity may be associated with periods of growth, with weight reduction, with pregnancy and with reversible illnesses.[3]

Diseases commonly result from complex causal webs rather than single factors.[4] Environmental pollution may contribute a number of strands to such webs. Other strands may arise from such diverse origins as genetic heritage, nutritional status and personal habits. Moreover, pollutant exposure may alter the severity of disease without altering its frequency.[3]

## Exposure Response Matrix

The duration and concentration of pollutant exposure are measures of the total dose to the human. It must be remembered that the rate at which the total dose is received may influence response.

Health effects of an environmental pollutant may be either short-lived (acute) or relatively permanent and irreversible (chronic). Acute or chronic effects may occur after a single exposure to a hazardous substance. This can be illustrated by acute radiation exposure which can cause acute radiation gastroenteritis and chronic leukemia. An air pollution episode may have similar effects though the permanent sequelae of acute episodes have never been adequately studied. Likewise, acute and chronic effects can result from long-term exposure. For example, excess acute respiratory illness and chronic respiratory disease have been repeatedly demonstrated in high air pollution exposure cities. Acute effects and short-term exposures are less difficult to study than chronic effects or long-term exposure. Moreover, the effects of dose rate, i.e., large dose in a short interval vs. repeated small doses over a long period, have seldom been investigated systematically. Little effort has been expended on the monitoring of long-term exposure and disentangling the causal webs underlying chronic effects of pollutant exposure. Such effort has been hampered, however, by methodologic difficulties.

## Evaluating Exposure-Response Relationships

Given adequate characterization of exposures and health effects, many additional considerations bear on the scientific validity of the exposure-response relationship. Hill has given an exposition of criteria that can be used to judge whether an observed exposure-disease relationship is causal. Hill's criteria were developed as guides for occupational health studies. With minimal modification, however, they can be applied to general population studies. These criteria are: Consistency of observed associations, coherence of results, plausibility of the association, and strength of the association. Brief descriptions of each of these as well as exposure response gradient, intervention, and control of covariates are included.

*1. Consistency of Observed Association:*

Consistency of observed association is, perhaps, the most important criterion. Does the health effect occur in various age, sex and race groups? Has the effect been repeatedly observed in different places, circumstances and times? Even small differences that are not quite statistically significant bear great weight in standard setting when the criterion of consistency is met. When the same effect is observed in a variety of population groups under varying conditions and at different times, the likelihood of a constant error or fallacy becomes progressively less.

*2. Coherence of Results:*

When animal studies, experimental human exposures, and epidemiologic data are coherent, i.e., they all demonstrate the same or similar health effects of exposure, the bits and pieces of evidence, when brought together, form a mosaic of health intelligence. By reinforcing results obtained in one research approach with studies in another, a coherent health intelligence system will provide scientifically strong and readily acceptable guides for costly air quality controls.

*3. Plausibility of the Association:*

Initial studies may uncover unexpected relationships between exposures and effects. Mere statistical associations must be rejected when no reasonable biological explanation, based on experimental evidence, can justify the association. On the other hand, when experimentation points to a disturbed physiologic process which may lead to some clinical manifestation, subsequent epidemiologic studies designed to test such hypotheses are well grounded in biologic plausibility. New exposure-response associations require support from other disciplines before causal inferences can be readily defended.

*4. Strength of the Association:*

When disease frequency is nine to tenfold greater in exposed than in non-exposed populations, the exposure-effect relationship is extremely strong. Relationships of this magnitude have been found in studies of cigarette smoking and respiratory disease including lung cancer and bronchitis. However, pollutant exposures are generally less intense and less reactive than cigarette smoke inhaled deeply into the lungs. Differences in exposure between high and low pollution neighborhoods seldom exceed factors of two or threefold in concentration. At present ambient air concentrations, relative differences in effects between high and low exposure areas are unlikely to be as striking as ratios observed in smokers vs. nonsmokers. However, for common and frequent disease events such as acute respiratory disease, relatively small differences in disease experience can have a large and costly impact.

*5. Exposure-Response Gradients:*

When a stepwise increase in exposure can be associated wit ha stepwise increase in the frequency of the adverse health effect, the evidence is strong for a cause-effect relationship. Linear relationships over an exposure gradient became increasingly difficult to explain by third intervening variables. Response gradients can be investigated in relation to exposure gradients across geographic areas, differences in length of residence in high exposure areas, and migration gradients constructed from various combinations of childhood and adult exposures of the same individuals. Human exposures occur naturally over an exposure gradient, especially over time and across areas. Exposure-response gradients obviously can be easily created in experimental settings.

*6. Intervention:*

Protection of health requires society to intervene in public exposures to air pollution. This intervention is largely based on observed exposure-health effects associations. As desirable air quality is achieved, will the frequency of adverse health effects be affected? If so, the causal nature of the exposure-response association is strongly supported. The high cost of air pollution control warrants a national program of community health and environmental surveillance in those places where achievement of required air quality will require vigorous abatement efforts.

*7. Control of Covariates:*

The causal nature of an exposure-response association is convincingly exposed when, after the effects of known covariates are first displayed, increased disease risk within covariate classes can be clearly demonstrated in high exposure populations. For example, in any studies of chronic bronchitis, prevalence, smokers and males show more disease than nonsmokers and females respectively. An air pollution-chronic bronchitis study should reveal the above smoking-sex differences in prevalence rates, thereby assuring readers that the study has internal consistency. If smoking-sex specific groups in high exposure neighborhoods have excess chronic bronchitis, the hypothesis that air pollution exposure causes excess chronic bronchitis is considerably more convincing than if the smoking-sex covariates were not analyzed. Most epidemiologic studies of air pollution require similar analysis of excess disease risk within covariate categories, with particular attention given to age, sex, smoking, socioeconomic level and duration of residence at current location. When covariates are systematically analyzed for relationships to the health indicator under study, residual excesses in disease frequency can be attributed to area differences in pollution exposure with a reasonable degree of confidence.[6]

## III. USE OF HEALTH RESEARCH DATA

The need for regulatory action is predicated on the demonstration of an effect at the levels of a pollutant to which the popular is or can be predicted to be exposed. The kinds of control options available to prevent such effects and to choose among are environmental standard setting, registration, certification, licensing, limiting use, banning and economic incentives such as effluent charges or taxes. For simplicity, the only option that will be address is environmental standard setting.

### Environmental Standard Setting

An environmental standard in itself is a complex issue. The standard is the maximum permissible level of exposure of the public to a pollutant permitted for a specified period of time. For a pollutant that is only present in the air, the environmental standard is the primary ambient air quality standard. However, when the pollutant is one where human exposure comes not only from the air, but from the food and water as well, the environmental standards become a set of standards, each representing an allocation of a portion of the permissible exposure through each of the media from which exposure comes. No such environmental standards have been set in the U.S.A. to date.[7]

The actual permissible level of exposure permitted must be decided in some philosophical framework. At the present time the framework varies somewhat by medium of exposure and by effect of exposure, in some instances by legislative mandate, and in others by differing beliefs of those responsible for the program. It certain is the scientist's role to be interested in these problems and to supply a coherent and rational basis for the framework within which regulatory decision-making can proceed.

Certain other concepts are useful to consider when discussing standard setting, such as threshold dose, dose-response, extrapolation of data from animal to man, and combined effects.

### Threshold Dose

A biological effect may not be observed until the dose reaches a certain level. This level is called the threshold dose and it may be defined as, the minimum dose required to produce a *detectable* effect. The concept is important from both a practical and a theoretical point of view, since a true threshold implies that below a given dose there is no adverse change whatsover with regard to toxic effect of the substance studied; this allows a safe limit or standard to be specified. As one uses more sensitive responses, thresholds decrease. For example, illness frequently is a more sensitive measure of response than death, and both will be preceded by physiologic changes heralding the onset of disease. For many pollutants, such as ionizing radiation, there may be no threshold for the response.[3]

The existence of a threshold may be due in part to the build-up in man of tolerance to the particular pollutant. Tolerance represents the ability of man to endure pollutant exposures without apparent ill effects. The level of tolerance to environmental agents may be directly related to a number of characteristics including age, sex, and nutritional state. The concept of adaptation signifies an increase in tolerance with long-term low-level exposure to a given adverse environment. Adaptation is characteristically related to the stressful components of the environment. The ability to adapt varies in a population and is determined by anatomic, physiologic and biochemical characteristics of individual organisms.

## "Safety"

Operationally the threshold concept is being used in standard setting activities in the United States regulatory agencies. However, many persons, finding that definitions for words such as "safe level," "no effect level" and "virtual safety" (among others) are being used to satisfy their intellectual recognition that such absolutes do not exist, are unhappy because the use of these terms gives the general public the implication of safety as an absolute. Thus, the concept of acceptable risk is being introduced and hopefully environmental agents soon will be assessed in terms of relative risk rather than as safe or unsafe.

## Dose Response

Dose-response data extracted from either community or experimental studies is crucial for the decisions that must be made. However, when the information is available, as it rarely is, it has been gathered in animals at rather high exposure levels and needs to be extrapolated to man who will be exposed at lower levels. Man, however, will generally be exposed in greater numbers and for longer periods of time so that low-level risk can be manifest. We urgently need more specific research into the issues of extrapolation so that better decisions can be made. Since dose-response information is frequently lacking, other techniques need to be considered.

## Extrapolation

Studies of laboratory animals have been used to assure that the toxicity risk from an environmental chemical is acceptably low. Such assurance should be given before a chemical is introduced into the environment in a fashion where there can be human exposure. However, relating animal studies to possible effects in man poses difficult and practical problems. As more chemicals for which there is sufficient toxicity testing are introduced into the environment where man can have similar exposure, we can learn more about the validity of the testing. Until that time we rely on such information as ". . . cancers in laboratory animals are essentially the same as human cancers and with a single possible exception, all known human carcinogens are carcinogens in laboratory animals, laboratory animal studies seem to predict for man. Thus, for ethical and practical reasons, data derived from use of laboratory animals for toxicity testing is the foundation of efforts to protect the public from the possible harmful effects of new and old chemicals in the environment."

The normal sequence of animal tests of a new chemical agent begins with studies to determine the mechanism through which laboratory animals respond to a compound, the nature of metabolites, the distribution of the parent compound, the metabolism in the tissues and organs, and the rates and routes of elimination of the compound or its derivatives. Comparison of the similarity of the results of these tests can be made among various animal species, thus laying the groundwork for prediction of the exposure events in man. Attempting to systematize this process is a prime area for research. Also, it must be remembered that this approach is most useful for observing effects that occur soon after the compound is administered, but is less useful for observing subtle long-term effects.

When attempting to bridge the gap between nonhuman and human toxicity one must recognize two steps involved in the extrapolation of laboratory toxicity data to man. First, the extrapolation of what might be called the average or median mouse to the average or median man. This consists of extrapolation from the average or even the most sensitive one, two or three laboratory animal species to an average man living under conditions with a commonly encountered genetic make-up. The second step requires identifying and allowing for variability in the human population. The median mouse to median man extrapolation is the easiest step. Biomedical research results are beginning to illuminate the differences and similarities between species and are beginning to provide a rational basis for extrapolation. The second step, allowing for variability and diversity—both genetic and environmental, is much harder.

There is, then, a basis for comparison of the median mouse to the median rat to the median dog to the median man. More and more patterns that are useful for extrapolation to man are being recognized and can be identified in the course of the study of the pharmacological disposition of a chemical. Most of the differences that have been observed suggest that man is more sensitive than the usual experimental animal, and this must be kept in mind when developing safety factors.

The following principles were stated in the Weinhouse and Rall[illegible] communication as this subject related to carcinogenesis testing:

1. The major principle that must be accepted if we are to deal at all with long-term assessments of toxicity in man is that effects in animals, properly qualified, are applicable to man.
2. Methodologies do not now exist to establish a safe threshold for long-term exposure to toxic agents.
3. The exposure of experimental animals to toxic agents in high doses is a necessary and valid method of assessing carcinogenetic hazard to man.
4. Agents should be assessed in terms of relative risks rather than as safe or unsafe.

## Combined Pollutant Effects

Physical-chemical interactions occur among the pollutants in the environment which in turn alter the biological activity of the contaminants as well as the reactions within the tissues of man and animals. The resulting bioeffects may be synergistic, additive or antagonistic, resulting in a reaction whose magnitude is greater, equal to, or less than the sum of the individual constituents.

One of the most recent and best demonstrations of synergism of two pollutants was presented by Bates and Hazucha in October, 1973.[illegible] Figure 3 shows the effect on four measures of pulmonary function of 0.37 ppm of $SO_2$ and 0.37 ppm of ozone exposure independently, and then the enhancement of effects when both gases are present simultaneously. It is interesting to note that the effect of the two gases administered simultaneously is greater than the sum of the effect of each individually; thus a synergistic effect has been demonstrated that might account for episodes in the U.S.A., Netherlands and Japan.

## Least Case and Worst Case Range Estimates[6]

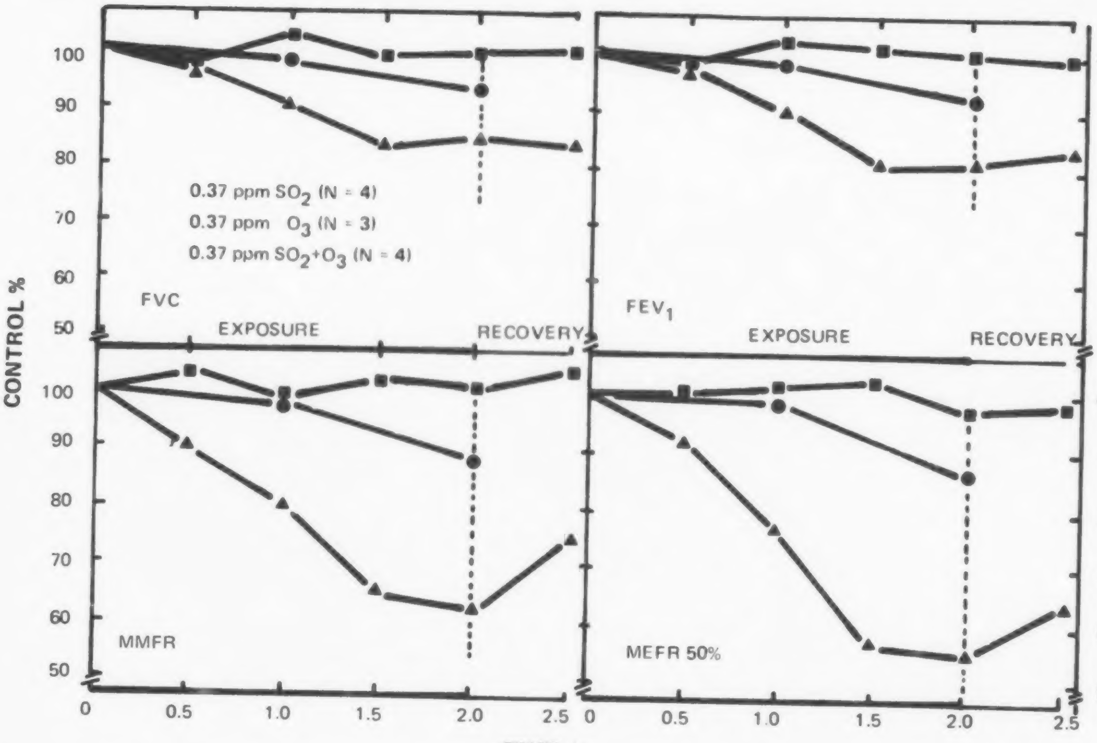
Derivation of range estimates based on least case and worst case assumptions is another way to develop a workable plan upon which action decisions can be based. In least case assumptions, quantitative assessment of exposure or response is made on the assumption that current exposures are not representative of long-term trends or that adverse health responses should first be attributed to all known covariates, and only residual excesses of illness frequency be attributed to pollutant exposure. For example, in the frequently-encountered situation where high exposure and low socioeconomic status geographically coincide, the effect of low economic level on illness frequency would be identified first. Any excess illness which could not be accounted for by economic level would be quantified as a residual effect. After all covariates were considered in turn, the final residual excess would be called an air pollution-related health effect. Least case estimates attribute the smallest possible effect to pollutant exposures, do not allow for interaction between covariates and exposure, and give a maximum quantitative estimate of human exposures associated with adverse responses.

Worst case assumptions attribute adverse health effects first only to covariates which are known to be strong determinants of disease frequency, such as cigarette smoking in relation to chronic bronchitis. But covariates which are not well founded as determinants of illness frequency are eliminated from final analysis, and air pollution exposure is assumed to have contributed to the relatively larger residual in excess illness frequency. Likewise, if information on past exposures is of low quality or unavailable, current exposures are assumed to represent past experience, and chronic disease frequencies are calculated as a function of exposures at current levels. Worst case assumptions, therefore, give minimum estimates of exposures associated with adverse health responses, and tend to maximize the proportion of disease frequency attributable to pollutant exposure.

The truth may lie at either end of the quantitative range which can be derived from least case and worse case assumptions. When health intelligence cannot give precise quantitative information, the decision maker should be provided with least and worst case range estimates. At this point, the degree of control becomes a function of other policy considerations, including control costs, alternate control strategies (and the health effects of these), the severity of magnitude of the effect, the population at risk, etc. Failure to present range estimates leaves less room for control options and forces decisions based on one set of numbers derived from arbitrary interpretations about study results.

## Least Cost Protective Standards[6]

Air quality criteria supply the existing rationale for standards. The process of standard setting requires policy decisions concerning protection of public health at least cost to society. Sound quantitative exposure response data allows the policy maker to make clear-cut decisions which can be defended.



**SOURCE:** **Proceedings of the Conference on Health Effects of Air Pollution, Oct. 3-5, 1973; Committee on Public Works, United States Senate 93-15, p. 534.**

**This figure shows the effect on four measures of pulmonary function of 0.37 parts per million of $SO_2$ and 0.37 parts per million of ozone independently, and then the enhancement of effect when both gases are present simultaneously.**

FIGURE 3. EFFECT OF INDEPENDENT AND SIMULTANEOUS EXPOSURES TO $SO_2$ AND $O_3$

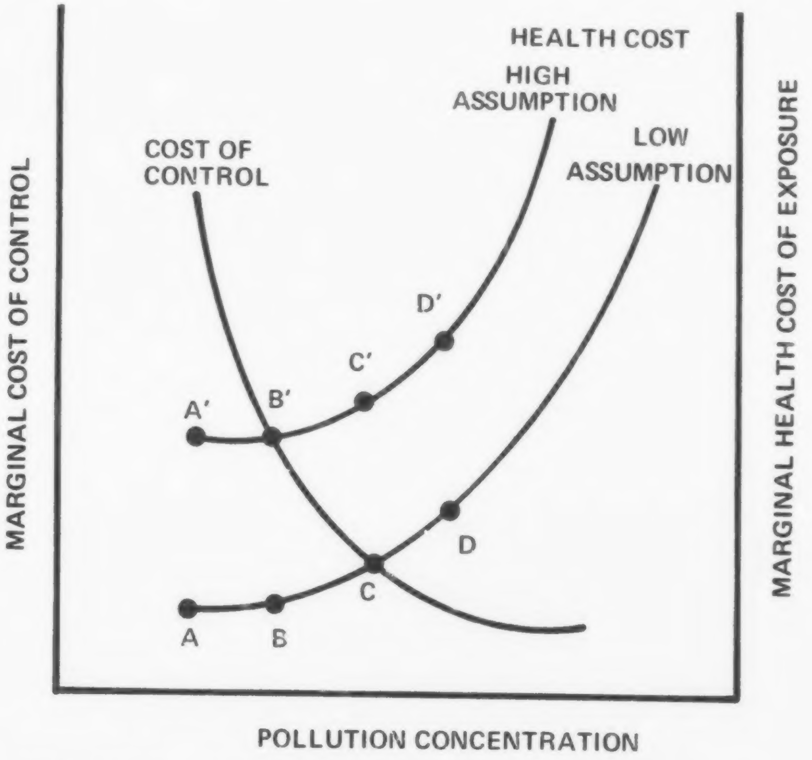
When range estimates are broad due to inadequate health intelligence one of two possible kinds of error may be made. The resulting standards could be too lenient and not protective enough or they could be too strict and excessively costly for the real benefit derived. Thus, society must pay a price for not generating adequate environmental health information. It must also be remembered that generation of health information itself can be costly and that in some instances it may be less costly to overcontrol than develop the specific information required.

Margins of safety built into standards are judgments required to bridge the gap between the inadequacy of health intelligence and the need to stop continued exposure to hazardous pollutant levels. Greater degrees of uncertainity generate larger safety margins. Unfortunately control costs usually increase exponentially as exposure is limited more severely. The health benefit from the additional degree of control may be insignificant or actually negative if overly stringent controls cause health disbenefits, as could occur if scarce resources are diverted from health care or protection to unnecessary pollution abatement.

The function of a health intelligence system in the standard setting process can be graphically displayed in the form of marginal cost curves (Figure 4). The true social cost of pollution is equal to the sum of the cost of control and the health and welfare cost of exposure. Health costs of disease, and more especially of physiological dysfunction, cannot be readily estimated with current scientific knowledge. Health and welfare costs increase with higher exposures, while control costs increase as exposures are lowered. True social costs of pollution are minimal where the two marginal cost curves intersect. If society will tolerate some pollutant-associated excess health costs, curve AD (Figure 4) may represent this social preference. On the other hand, if society places a very high premium on preventable disease or reducing risk of disease, curve A'D' (Figure 4) may well represent this attitude. On both curves, the points B or B' represent the lowest pollutant exposure associated with adverse health (or welfare) effects. Under the low health cost assumption (curve AD), society will accept the health costs (and associated adverse health effects) represented by the vertical distance between B and C. Under the high cost assumption, society will not allow any adverse health effect caused by exposure and will require control to point B', where high marginal cost of control is equated with derived health benefit.

Excessively stringent standards are shown at points A and B under the low cost assumption and at point A' under the high assumption, while too lenient standards are reflected in point D under the low assumption and C' and D' under the high assumption. The least cost standard, in terms of true social costs, is at point C under the low health cost assumption and at point B' under the high assumption. Least cost standards protective of health require adequate quantitative exposure-response information, knowledge of control costs and range estimates of health cost of pollution effects.

Even range estimates of pollutant related health costs are difficult to derive. Many direct and indirect health costs must be considered, including: Immediate and delayed health effects of short-term and long-term exposure; the contribution of pollutant exposure to the occurrence and severity of major public health problems such as acute and chronic respiratory diseases, heart disease, cancer and congenital defects; and adverse health effects of control strategies. While these estimates are formidable tasks, failure to make the estimates will prevent even the possibility of achieving least social cost standards, except entirely fortuitously. Required annual reports of air pollution control costs clearly warrant better and more quantitative health effects data and



LEAST COST STANDARD: C OR B'
EXCESSIVELY STRINGENT STANDARD: A, B, A'
EXCESSIVELY LENIENT STANDARD: D, C', D'

FIGURE 4. LEAST COST PROTECTIVE STANDARDS

concerted efforts to convert health effects to health costs.

Recently the National Academy of Sciences was asked to review the health basis for the ambient air quality standards. One part of that review was a Conference on Health Effects of Air Pollutants that was held on October 3-5, 1973.[10] All existing U.S.A. ambient air quality data were presented as least case, worst case and best judgment values.

## BIBLIOGRAPHY

1. WHO: Air Quality Criteria and Guide for Urban Air Pollutants. Wld. Hlth. Org. Techn. Rep. Ser. 1972, No. 506.
2. Newill, V. A.: PSAC Briefing on the EPA Air Pollution Health Effects Research Program, Washington, D.C. May 22, 1972. This briefing was based on many in-house reports from the Division of Health Effects Research, NERC-RTP, EPA, North Carolina.
3. Message from the President of the United States Transmitting the Report of the Department of Health, Education and Welfare and the Environmental Protection Agency on the Health Effects of Environmental Pollution, Pursuant to Title V of Public Law 91-515. 92nd Congress, House of Representatives Document No. 92-241, February 1, 1972.
4. MacMahon, B.; Pugh, T. F.; and Ipsen, J.: Epidemiologic Methods. Little, Brown and Company, Boston, Massachusetts, 1960, pp. 18-21.
5. Hill, A. B.: The environment and disease: Association or causation. Proc. Royal Soc. Med. 58:295-300, 1965.
6. Shy, C. M.: Health intelligence for air quality standards. Presented at the Meeting of the President's Air Quality Advisory Board, St. Louis, Missouri, March 27, 1973.
7. Newill, V. A.: Nature and source of pollutants. Presented at the American Academy of Pediatrics Conference, Evanston, Illinois, June 11, 1973.
8. Weinhouse, S. and Rall, D.: Personal communication.
9. Bates, D. V. and Hazucha, M.: The short-term effects of ozone on the human lung. Proceedings of the Conference on Health Effects of Air Pollutants, NAS, NRC, October 3-5, 1973, pp. 513-540. U.S. Congress Document Serial No. 93-15, November 1973.
10. Finklea, J.: Conceptual basis for establishing standards. Proceedings of the Conference on Health Effects of Air Pollutants, NAS, NRC, October 3-5, 1973, pp. 619-668. U.S. Congress Document Serial No. 93-15, November 1973.

# Maximum Permissible Human Stress Pollutants with Multiple Pathways: Ionizing Radiation, Asbestos and Lead

John H. Knelson *

## I. INTRODUCTION

All human activity is associated with a degree of stress; that is with a degree of risk to the health and welfare of the human organism. Only recently, however, has man begun to contemplate seriously the degree of stress and the risks associated with rapidly evolving technology. Initially, these concerns were focused directly at the most obvious site of risk, the workplace itself. Now, concepts from industrial hygiene and medicine are serving as important bases for development of the more general field of environmental hygiene. As technology becomes more complex and pervasive, the problems of environmental hygiene also become more complex and more pervasive; obviously then, regulatory decisions based on considerations of environmental health have increasingly significant social consequences.

Assessment of the influence of environmental factors on human health is one of the most important components of comprehensive environmental analysis. The measurement of human health, however, is in itself a complex and incompletely understood task, and it becomes increasingly complex when one considers the multiplicity of environmental stressors which may act upon the body. It is essential, therefore, that some simplifying concepts be employed to organize and interpret environmental health effects data. How do we define "Maximum Permissible Human Stress"? How do we determine environmental standards for pollutants that permeate the environment, enter the human body, and exert influence through multiple mechanisms? These are parts of the bigger question, "How should we modify or redirect our technological development to assure that environmental hazards are maintained at a level that represents the best balance between acceptable risk and acceptable cost?"

## II. DOSE-RESPONSE RELATIONSHIPS

The determination of an association between an environmental agent and a putative health effect requires that the exposure concentration in the environment be measured accurately and unambiguously. This rather self-evident prerequisite to health effects studies is frequently not met. Although conceptually simple, it is in practice often very difficult to measure with acceptable accuracy the pollutant dose to which a population is exposed. This problem is compounded when the agent exists in several forms in the environment and enters the human organism through a variety of pathways and from a number of sources. Even when good chemical analysis is possible, determination of the actual dose delivered to the individual is difficult. Environmental monitoring at a particular point gives a relatively poor estimate of the spatial variability of pollutant concentrations within an area—and the exposure over time of an individual residing in that area is even more indeterminate.

Some materials, such as lead, accumulate in human tissue. That accumulation, a function of dosage integrated over some period of time, can be expressed as "pollutant body burden." This total body burden, the total quantity of a substance distributed within body tissues or physiologic compartments, cannot be partitioned into contributions derived from each of the multiple portals of entry, nor does body burden correlate consistently with resultant health effects.

In the case of asbestos, monitoring of exposure becomes excruciatingly difficult for not only is particle composition crucial, but size and aerodynamic characteristics are also important. Although "ferruginous bodies" are found in the sputum of asbestos workers, no way has been developed to associate these bodies quantitatively with the intensity of environmental exposure. With ionizing radiation, convenient personal monitors—film badges—are available. But they do not measure that fraction of total body dose inhaled or ingested as radionuclides. In general, then, measurement of the dose of any environmental pollutant an individual receives is still technically difficult.

---

*Acting Director, Human Studies Laboratory, National Environmental Protection Agency, U.S. Environmental Protection Agency

Measurement of health effects attributable to environmental pollution is also fraught with many methodologic difficulties. Definition of health and description of statistically significant deviation from the norm is greatly complicated by the multitude of factors that determine the health status of any population. These difficulties in the measurement of health effects pervert the meaning of threshold. Though paying homage to the true definition, in the real world we set a threshold level at the point where available health effects data fail to show any significant dose-response relationships. Historically, the more data accumulated the lower the threshold level is set. In this context, the threshold, when not just a statistical artifact, is purely an operational concept which contains more information concerning the level at which our uncertainty begins than about the underlying reality.

As the tools for measuring dose as well as those for measuring human response are improved, it becomes possible to describe a continuum of effect down toward zero dose. Thus, even a very small environmental insult probably has a very small but measurable impact on human health. It is then no longer a question of "What is a safe level?" Rather, the question becomes "What level of risk is acceptable?" The concept of threshold limit value is replaced with the much more complicated concepts of Cost-Benefit analysis, which allow us to define Acceptable Risks. It is no longer sufficient for the scientist to provide a threshold value and suggest a margin of safety. He must provide the decisionmakers with the best possible objective description of health effects as related to dose, and also with an estimate of his degree of uncertainty concerning that relationship.

## III. ENVIRONMENTAL COSTRESSORS

Implicit in the description of Dose-Response relationships is the ranking by relative priority of the described relationships. A comprehensive control strategy requires knowledge not only of the well described health effects of a particular environmental agent, but knowledge of the importance of that agent relative to the importance of other risk factors, not all environmental, which may lead to the same health effects. The risk attributable to a particular environmental agent must be judged in comparison with the risks attributable to other factors, such as with eating, drinking, smoking, physical activity, and mode of life in general. Only in this way may we efficiently employ society's finite resources; the cost-benefit ratio should be minimized not just with respect to all activities which come under the broad umbrella of Public Health. It should be clear, however, that control activities will not invariably be directed at the most important risk factors, as risk factors are not all equally amenable to control.

Risk factors not only sum up, they may also interact. It is well known that smoking increases the risk of industrial asbestosis. Ionizing radiation in some forms, smoking, and asbestos fiber inhalation are all associated with various types of lung cancer. Yet, we are almost totally ignorant of the relative contribution of these factors, much less the possible interactions among the three. More recently, we have begun to suspect that low level lead intoxication and certain kinds of ionizing radiation may act synergistically to produce some forms of mental retardation. We are only beginning to explore the relative importance of environmental costressors.

## IV. RANGE OF SUSCEPTIBILITY

Humans are not uniformly susceptible to environmental stress. The historic episodes that first drew worldwide attention to the lethal consequences of severe air pollution affected primarily the very young, the aged, and the infirm. Those early observations were evident even to the untrained. Since then, more careful documentation of range of susceptibility has been possible.

In epidemiologic studies, the strongest correlations with pollutant exposure have been seen in the incidence of respiratory infection in children, in the increased number of attacks in asthmatics, and in the increased intensity of symptoms in people with cardiopulmonary disease.

Ionizing radiation is known to have an especially deleterious effect on the developing mammal, but we have no firm evidence that any subsets of the adult population are particularly susceptible. Excessive lead intake, by whatever pathway, is associated with impaired central nervous system development in young mammals, and may be particularly harmful in people of all ages with some types of kidney disease. Particular susceptibility to inhaled or ingested asbestos in the absence of other environmental costressors has not been described up to now.

## V. SPECTRUM OF RESPONSE

Whereas there is a range of susceptibility within a population, "spectrum of response" refers to the various ways an individual can manifest the effects of environmental stress. The spectrum of response is usually a function of pollutant concentration, duration of exposure and the presence of costressors such as exercise, thermal stress, or other coexisting pollutants. We have somewhat arbitrarily classified the spectrum into categories ranging from the mildest to the most severe. They are: (1) pollutant body burden, (2) changes of uncertain significance, (3) pathological changes, (4) morbidity, and (5) mortality. Because the proportion of any population falling into any category is greatest with (1) and the least with (5), it is evident that the concept of "range of susceptibility" is closely related to that of "spectrum of response." Combining the concepts of dose, spectrum of response, and range of susceptibility allows us to construct a three-dimensional dose-response surface from which both risk and the proportion of the population at risk can be estimated for each dosage level. This kind of prediction is the final goal of environmental health effects research. We are trying to answer the question, "How many individuals in each category of our population are subjected to what levels of risk when subjected to pollutant concentrations within a given range?"

## VI. THE EXPOSURE-RESPONSE MATRIX

Comprehensive environmental analysis as a foundation for comprehensive environmental control requires not only a firm health effects data base, but interpretation of the data in a manner useful to those responsible for establishing the control strategy. The concepts necessary for that interpretation have been described. Assembling the proper variables into an exposure-response matrix should provide, for each environmental stressor, a point of departure for eventually assessing the Maximum Permissible Human Stress attributable to pollutants with mulitple pathways.

The exposure-response matrix uses dose, established by environmental monitoring and modeling or by pollutant burden analysis, as the independent variable. The dependent variable is the Spectrum of Response. To these are added an intervening variable, the Range of Susceptibility. Without this third covariate, such a plot is simply the classic dose/response relationship. With the "z" axis representing the Range of Susceptibility, we have additional information of use in total assessment of risk. It allows us to relate the degree of response at a particular dose to specifically susceptible population subsets.

Estimates of Maximum Permissible Human Stress can be derived from the exposure-response matrix in combination with a sector of the relevant population sizes for each population subset. From the set of outcomes, a weighted total must be generated, the set or sector of weights expressing the importance of the outcome. These weights are necessarily judgmental. However, their use with the exposure-response matrix does at least provide a tool for applying judgmental decisions objectively and uniformly in developing the cost/benefit information that is needed for establishing environmental standards.

## VII. ENVIRONMENTAL AGENTS WITH MULTIPLE PATHWAYS: IONIZING RADIATION, ASBESTOS AND LEAD

Three examples of environmental agents with multiple pathways for entry into the human and multiple mechanisms of action will be discussed briefly to elaborate some of the principles for establishing maximum permissible human stress. These examples have been chosen primarily because they illustrate the problems involved in dealing with pollutants of multiple pathways, not because the problems with these pollutants have been solved. It is anticipated, however, that analyses of these problems will demonstrate some helpful first steps in arriving at answers when one is dealing with comprehensive environmental analysis and control.

### Ionizing Radiation

This example of an environmental stressor, or category of stressors, is particularly important because it is probably the one for which we have more health effects data than any other. Thus, some important analytic, conceptual, and control philosophy precedents have been set. Many prototype procedures for environmental analysis and control have been developed by a generation of scientists who devoted their careers to studying the problems of ionizing radiation. It is an environmental stressor of multiple pathways *par excellence* for it is transported through all media of the biosphere, entering the human via all imaginable portals.

It is significant to note that the most recent summary of effects on populations exposed to environmental ionizing radiation begins with a careful comprehensive analysis of nature, sources, and transport mechanisms of the environmental stressor. Without this analysis, estimation of health effects is inadequate. Too frequently, toxicologic evaluation of other environmental stressors is attempted without sufficient data concerning pathways from the various sources to the human recipient.

Biologic effects of ionizing irradiation are classified according to influence on genetic material, growth and development, and somatic health. Genetic effects predominated early concern, and have become increasingly important as more has been learned of the molecular basis for disease. Increased knowledge of genetic effects has not lead to a greater precision of dose-response estimates, however, as new complexities have been discovered. The principle of dose-rate independence has lost some credibility as mechanisms for cellular repair of radiation damage become better understood. However, dose-rate delivery during an organism's relatively short developmental period must still be considered in establishing life-span standards because of the greater susceptibility of the developing organism to effects of ionizing radiation. Important analogies with other environmental stressors emerge as one considers the problems of establishing other dose-rate standards. Averaging times for most pollutant standards are established with crude dose-rate effects data, if any.

Many environmental stresses have more pronounced effects during developmental stages of the organism, as does ionizing radiation, and this serves not only to emphasize the range of susceptibility concept, but also to raise the problem of pollutant interactions. Other environmental agents, especially heavy metals, drugs, and food additives are known to be preferentially distributed to certain tissues during the development period and these may exert a synergistic effect with other environmental factors which must be recognized in a comprehensive analysis of maximum permissible stress.

The somatic effects of ionizing irradiation, e.g., cancer induction, provide, perhaps, the most striking example that multiple pathways and interactions of environmental stressors must be considered in establishing acceptable risk levels for each pollutant. Many pollutants are carcinogenic. They permeate the environment, entering the host by all portals. Their effects may well be additive or even synergistic. They represent the most important challenge to comprehensive environmental analysis.

The comprehensive environmental analysis of ionizing radiation has been more extensive than for any other environmental stress. Yet, even with the extensive data available, establishing the exposure/response matrix as described is very difficult. Estimates of acceptable risk are possible only within a range of one order of magnitude. Although genetic, developmental, and somatic effects have been explored, the genetic effects of radiation are most important in estimating acceptable risk.

Four techniques have been used to calculate genetic risk:

(1) Risk relative to the effects of natural background radiation;
(2) Risk for specific genetic conditions;
(3) Risk for severe malformation and disease;
(4) Risk in terms of overall ill health.

Natural background radiation averages about 10 mrem per year. However, there is considerable variation in this figure depending on geographic location, type of dwelling, and altitude. Varying concentrations of naturally occurring minerals that emit radiation account for most of the geographic differences. Structures built of stone containing these minerals will subject the inhabitants to more background radiation than wooden buildings. Screening of cosmic (mostly solar) radiation by the atmosphere decreases with altitude. Thus, man throughout his existence has been exposed to radiation, and the effect of that natural or background level serves as a convenient denominator for expressing relative risk. If man accepts a range of risk, depending on the factors described, from background radiation approaching zero to that over 200 mrem per year, it can be argued that he should accept additional unspecified risk the same order of magnitude associated with environmental stressors or man's creation. "If the genetically equivalent significant exposture is kept well below this [background] amount, we are assured that the additional consequences will neither differ in kind from those which we have experienced throughout human history nor exceed them in quantity" (BEIR Report, 1972). This estimate, based only on an estimate of dose and none of health effect is more philosophical than scientific, but the advantage of avoiding all the assumptions of human radiation genetics.

Risk estimates for specific genetic abnormalities can be made by comparing the current incidence of observed genetic conditions to those expected from increased radiation. There are no directly applicable human data, so those from experiments with mice must be used. Chronic radiation to developing mouse sperm cells results in about $0.5 \times 10^{-7}$ recessive mutations per rem per gene. Developing oocytes are much more resistant. Therefore, the average for both sexes is less than $10^{-7}$, and must because of the uncertainties in the data be estimated as somewhere between $10^{-8}$ and $10^{-7}$ per rem per gene.

Spontaneous mutation rates in the human are estimated to be less than $10^{-6}$. This is 10 to 100 times the rate induced per rem. Thus, the ratio of radiation-induced to naturally occurring mutations is 1/10 to 1/100, and the additional dosage necessary to double the spontaneous mutation rate is 10 to 100 rem. This is strictly an "order of magnitude" estimate. More subtle arithmetic might place the estimate at 20 to 200 rem for a "doubling dose."

Another statistical approach to the concept of "doubling dose" utilizes data from Hiroshima and Nagasaki. The observed death rates in children of exposed survivors does not differ from matched population controls who were not exposed. However, if some assumptions concerning average acute parental dose (100 rem each) and liveborn death rate from "spontaneous" genetic abnormalities are accepted, the chances are less than 0.10 that the true death rate was as great as one percent higher in children of irradiated parents. Thus, 100 rem at most resulted in an effect twice that occurring without additional radiation, and the "doubling dose" would be 50 rem. This is within the estimate of 20 to 200 rem for "doubling dose" made on the basis of entirely different assumptions.

To estimate the impact of induced genetic impairment on the overall risk of severe malformation and disease, one must know the increase in induced genetic impairment and the contribution of such impairment to disease in general. Using the estimate of 20–200 rem as the "doubling dose," the increase in relative mutation risk is in the range of 0.005–0.05 per rem. The continuous exposure rate of 5 rem per 30 years (or 170 mrem per year) would, in equilibrium, cause an increase of 2.5–25 percent in the overall incidence of genetically-related disease. But what fraction of malformations and other recognizable disease states are mutation-related? Estimates of the total burden of human disease related to hereditary factors are virtually meaningless. Therefore, even the best estimates of radiation-related genetic abnormalities (or spectrum of response) mean little until better data concerning the range of susceptibility within the population are available.

Risks in terms of overall ill health are included because the other estimates deal with more specific disease states that could result in an under-estimate of radiation risk to health in general. Again, only order of magnitude relative risk estimates are possible with available data. The most reasonable estimate is that 20 percent of all ill health may have some genetic component. If 20 rem is the mutation rate doubling dose, an exposure of 5 rem per generation (5 rem per 30 years or 170 mrem per year) would increase the total ill health burden over a long period of time by 5 percent. If 200 rem is the doubling dose, the increased burden would be 0.5 percent.

Thus, with a variety of assumptions and with a relatively large data base, the standard for the multiple pathway environmental stressor—ionizing radiation—is set at 170 mrem per year. This standard has been established without the refinement of the three-dimensional exposure-response matrix because of the many difficulties cited. However, many attempts have been made to provide the data necessary. The inadequacies are logistic, not conceptual. As more is learned of ionizing radiation effects, the matrix can eventually be completed to provide a more accurate estimate of maximum permissible human stress.

## Asbestos

Although associated historically only with occupational exposure, over the past ten years considerable evidence has accumulated which associated asbestos disease with nonoccupational exposure as well. This evidence also suggests a spectrum of response ranging from those observed in persons exposed in their occupations to those observed in persons only indirectly exposed by family contact or by neighborhood location.

Factors which complicate studies of the effects of asbestos air pollution include:

(1) The long lapse period between exposure and illness (often 20, 30 or 40 years). The long periods between exposure to asbestos and development of mesotheliomas that have been observed could be indicative of a breakdown of human defenses in elderly people. This suggests a range of susceptibility, i.e., individuals past the age of asbestos workers who live in the vicinity of emittor industries may be affected by acute or chronic exposure to relatively low levels of asbestos.

(2) Lack of knowledge concerning the significance of intensity of exposure. Although there is evidence suggtsting a continuous spectrum of response for lung disease, there are no data for establishing dose-response relationships for the development of cancer. This dilemma emphasizes the difficulty inherent in establishing standards for a substance of multiple pathways and different pathophysiologic mechanisms.

(3) Size and variety of asbestos fibers. Whereas the occupational exposure limit is based on counts of asbestos fibers $> 5\mu$ in length, this counting technique has little usefulness in ambient situations. Fibers emitted into the ambient air break up into fibrils the detection of which requires the utilization of electron microscopy. Consequently, equivalent quantities of asbestos pollution cannot be detected in occupational and ambient situations by determinations of numbers of particles $> 5\mu$ in length. When asbestos is implanted into the pleura of rats, cancer is produced only by fibers greater than $5\mu$ in length, and usually at least $40\mu$ in length. Fibers less than $5\mu$ in length appear to be less carcinogenic. However, this is not interpreted as indicating that shorter fibers are completely harmless. Preliminary studies of asbestos pollution in water indicate that by far the greater portion of the asbestos consists of small particles, although some larger fibers do occur. Hence, the problem of asbestos pollution in water may be quite different from the problem of airborne asbestos pollution.

(4) The synergistic effects of cigarette smoking or exposure to other concomitant air pollutants. In industry, insulation workers who smoked had eight times the risk of lung cancer of smoking workers exposed to asbestos and 90 times the risk of nonsmoking workers not exposed to asbestos. Comparable increased risks may be experienced by smokers who live in the vicinity of asbestos industries.

Since timely and accurate measurements of total quantities of asbestos in the ambient air would be impractical, or impossible, in most areas, protection of the general population must rely on control of emissions. This heavy reliance on emission control is similar to the strategy applied for ionizing radiation. Considering the complete inability to quantitate the effects related to specific exposure, plus the repeated association of fatal disease with nonoccupational exposure, it would appear desirable that emissions be restricted to the limits compatible with the best available control techniques. Such control recommendations should not be construed to be recommending elimination of all asbestos from the ambient air. Nonoccupational illness has been associated almost exclusively with indirect occupational exposure or residence in the immediate vicinity of an asbestos source. There is no evidence that the small number of fibers to which the general population is exposed affects health or longevity; this is similar to the situation with "background" ionizing radiation.

With available data, how much of the exposure-response matrix can be filled in for asbestos? The first problem is environmental monitoring. Asbestos, a hydrated mineral silicate, occurs in several forms. Most widely used in United States' industry is chrysotile, a magnesium silicate. Other forms include iron, sodium, and calcium as well as magnesium silicate. Virtualy nothing is known of the relative toxicity of these forms of asbestos from data pertaining to human exposure, and very little comparative animal toxicology is available. Because of this, monitoring virtually never includes specific analysis of composition. Instead, the airborne asbestos filaments are collected on impactor plates or filters, and the particles sized and counted by microscopy. The relative toxicity of longer fibers (greater than 5 micrometers) and motes (nonfibrous particles) is a question of continuing concern. Although some studies do show a deleterious effect of motes, their toxicity is apparently much lower than that of the longer fibers. Whereas pulmonary pneumoconiosis and silicosis due to particulate silicate inhalation have much in common with pulmonary asbestosis, it seems only the fibrous form of asbestos is associated with gastrointestinal carcinoma and mesothelioma. With respect to airborne pathways, operational monitoring systems as well as particle or filament sizing and counting techniques provide less than adequate data. These may be improved as automated optical devices are developed and qualitative determinations using such techniques as microprobe analysis are employed. With respect to ingestion, especially from drinking water that has been transported through asbestos pipes, the monitoring problem is even more complicated. Again, the source of asbestos fibers associated with gastrointestinal carcinoma and mesothelioma is ambiguous. It is difficult to know how much of the ingested dose is first inhaled and then swallowed after tracheobronchial mucocilliary transport, and how much may come directly from contaminated drinking water. The only way this question can be answered for exposed human populations is with much more sophisticated and comprehensive environmental monitoring. One can appreciate, therefore, the difficulties in characterizing environmental exposure to asbestos for the purpose of establishing the exposure-response matrix.

The spectrum of response for asbestos exposure has been mentioned. In all cases, however, the health effect is disabling and eventually fatal for some individuals. Therefore, all effects lie at the extreme end of the response spectrum, and large weighting factors are applicable.

Range of susceptibility estimates are made difficult by the inadequate assessment of population at risk and poor understanding of the importance of environmental costressors. Tobacco smoking seems the most obviously implicated costressor. But absorption of carcinogenic hydrocarbons with subsequent desorption after inhalation has also been suggested as an additional cofactor. Trace element contamination of the mineral matrix may also play a role in the overall toxicologic assessment of asbestos. In general, the assumption has been made that effects of asbestos exposure are dose-rate independent and the concept of "fiber years per cubic centimeter of air" has been employed. From estimates of population risk, the probability of early signs of lung disease is less than one percent for 100 fibers years per cc. This would be equivalent to four fibers per cc for 25 years or two fibers per cc for 50 years exposure. There are virtually no data to estimate size of population at risk, or range of susceptibility, for asbestos-related gastrointestinal carcinoma and mesothelioma. Thus, as with ionizing radiation, standards must be set as low as possible because of lack of health effects data and necessity for maintaining a reasonable and prudent safety margin. A great deal more information is necessary before the exposure-response matrix for all pathways and effects of asbestos can be substantially filled out.

## Lead

Environmental standards for lead which have been proposed, take into account more information on spectrum of response and range of susceptibility than is available for most other pollutants of multiple pathways. There is little controversy over the basic toxicology of lead compounds. Toxicity of lead has been recognized since antiquity. It has most usually been associated with oral ingestion, however, and only more recently has airborne lead assumed greater importance. The concern with lead as a pollutant of multiple pathways has emerged since the introduction of organic lead compounds to automotive vehicle fuel because of their antiknock properties. There has been some controversy over the relative importance of inhaled versus ingested lead.

As usual, environmental monitoring has necessarily preceded the design of good health effects studies. There is now little argument that the air of urbanized regions has higher lead concentrations than that of rural areas. Tentative results indicate an amelioration of air quality in urban areas where significant decreases in automotive lead emissions due to reduction of lead in gasoline have occurred. Finally, the contribution of airborne lead to ingested lead via dustfall merits comment. As with ionizing radiation and asbestos, lead represents a complex environmental pollutant of several related pathways.

The spectrum of response for lead, unlike that for ionizing radiation or asbestos, has been studied from (1) pollutant body burden through (5) mortality. Pollutant body burden is usually expressed as blood lead concentration. In addition, however, considerable animal data are available to describe lead distribution in various body tissues and compartments. The validity of extrapolation of these data to the human is still not completely certain, but some corroborating information from analysis of autopsy tissues is becoming available. A change of uncertain physiologic significance most clearly related to increased lead body burden is found in decreased activity of red blood cell aminolevulinic acid dehydrase. This enzyme is associated with hemoglobin synthesis, but is also present in other tissues such as the central nervous system where its function is not understood. Changes in activity of this enzyme correlate well with changes in blood lead concentrations, but this is of uncertain physiologic significance because lowered enzyme levels are not unambiguously related to any specific diesase state. Changes in renal function associated with higher exposures represent a specific precursor of morbidity. These changse are usually associated with industrial exposure or with lead ingestion. Animal studies, however, suggest that very low exposure to lead results in observable changes in renal microstructure, and that there is a dose-response relationship continuing down toward zero with no demonstrable threshold. Moreover, there is evidence suggesting that repeated low-level exposure may alter pathways of detoxification and excretion without any signs of toxicity. Finally, of course the overt morbidity and mortality of central nervous system and renal disease related to lead intoxication is well-known. Less well appreciated or understood at this time is the possible subtle interaction of lead with many other environmental factors, including ionizing radiation, to produce mental retardation in children. One sees for lead, therefore, a wide spectrum of response irrespective of the pathway into the human.

Whereas there has been fairly good understanding of the relationship between dietary lead and resulting blood lead or pollutant burden level, only recently have reasonably good data for respiratory lead uptake become available. Studies exposing men to carefully controlled low air lead provided data for a rather well-defined respiratory lead uptake model. Population studies have resulted in comparable estimates of the relationship of air lead to total body pollutant burden.

Females residing in urban area and exposed to higher air lead levels were found to have significantly higher blood lead levels than those living in suburban areas where air lead levels were lower. In Philadelphia, 11 percent of women living in the urban area had blood lead levels of 29 $\mu$g/100 gm or greater. The average urban air lead concentration there was 1.67 $\mu$g/m$^3$ during the period of the study. Similar findings in Chicago were: 3.4 percent of the urban women studied had blood lead levels of 29 $\mu$g/100 gm or greater with an average air lead of 1.76 $\mu$g/m$^3$. In New York, however, only 1.4 percent equaled or exceeded the 29$\mu$g/100 gm level at mean air concentration of 2.08 $\mu$g/m$^3$. Extrapolating these percentages to the entire population base of these urban areas would, of course, involve a large number of people with blood levels that can be considered excessive.

Another recent study has found significantly increased blood lead levels in women living near a well-traveled roadway compared to those living further away. Air lead exposures were particularly well-quantitated in this study because measurements were made on the front porches and inside the homes where the women lived. The average air lead just outside the homes nearest the highway was 4.6 $\mu$g/m$^3$ and that inside the home was 2.3 $\mu$g/m$^3$. The air lead levels just outside and inside the homes 122 meters (400 feet) from the highway were 2.24 $\mu$g/m$^3$ and 1.57 $\mu$g/m$^3$, respectively.

Epidemiologic evidence indicates that while mean blood lead levels in populations are not significantly elevated by exposure to aid lead of about 2 $\mu$g/m$^3$, the frequency distribution of these blood lead values do differ so that significant urban-suburban differences do exist, and large numbers of people do have blood levels equaling or exceeding 30 $\mu$g/100 ml level at average air lead exposure levels of 2$\mu$g/m$^3$ and less.

With respect to range of susceptibility, in the case of lead it is becoming evident that children, the fetus and newborn in particular, constitute a specially susceptible group. This observation is certainly related to the increased incidence of overt lead poisoning in children due to their eating lead-based paint. Not equally appreciated, however, is the importance of their eating lead-containing dust and dirt from atmospheric fall-out. Recent evidence indicates that children may absorb considerably more of their oral lead intake than do adults. Although opinion is far from unanimous on the mechanisms of lead absorption from the respiratory tract, a significant fraction of that inhaled and deposited there may eventually find its way to the gastrointestinal tract via mucociliary tracheobronchial clearance. Greater enteric absorption of lead by children is thus important from several aspects. Finally, the respiratory dose per unit body weight may well be greater for children than adults because their level of ventilation is greater at rest and their level of activity is higher. There is no evidence that their detoxification and excretory mechanisms for lead are proportionally accelerated. Neither is there any to suggest it is not. Taking into account reported overt clinical toxicity in children of lead at blood levels of 40—50 μg/100 ml, the other considerations discussed above, and the most minimal of safety margins, an acceptable blood lead concentration for children of 30μg/100 ml seems prudent and reasonable. Because of the reported high correlation between newborn and maternal blood levels, expectant mothers should also be protected from achieving blood lead levels of 30 μg/100 ml or higher.

It can be concluded, therefore, that a prudent and reasonable upper limit for blood lead concentration, one with a certain margin of safety, that can be used as a criterion for setting an air quality standard is 30 μg/100 ml. This level applies to a large fraction of the population; that is, newborns, children and expectant mothers. Adults may tolerate, with a margin of safety, levels somewhat higher than this. Because it is not feasible to implement different ambient air quality standards for children and expectant mothers as opposed to those for the rest of the population, the lower value must be chosen.

The exposure-response matrix for lead can be completed with a degree of confidence not yet possible for most pollutants. Imperfect as it is, blood lead concentration nevertheless serves as an important index of integrated lead exposure from different pathways over relatively long periods of time. Comprehensive environmental standards for air, food and water can be established to maintain the quality of the environment above that related to acceptable risk.

## VIII. CONCLUSIONS

Assessing the influence of environmental agents on public health is a difficult multifactorial problem. Composition and concentration, sources, distribution, and bioavailability of pollutants must be determined before meaningful estimates of maximum permissible human stress or acceptable risk are possible. As comprehensive environmental analysis gives improved estimates of existing dose and dose-rate factors, biological effects data become more relevant. Spectrum of response and range of susceptibility provide the tools for establishing an exposure-response matrix from which some objective cost/benefit analyses can be drawn. The problem of assessing combined effects due to pollutants with multiple pathways has barely begun. It will be best pursued using firm principles and precedents established in evaluating single pollutant effects.

# Complex Analysis of the Environment, Approaches to a Determination of the Permissible Loads on the Natural Environment and to a Justification for Monitoring

Yu. A. Izrael [*]

Two interrelated questions arise, given the extremely varied anthropogenic effect on the natural environment:

(1) What is the state of the biosphere today (on a global, ecosystem, regional, or local scale); is this a crisis state?

(2) What is the possible, permissible load of the different effects, taken in the existing complex, on the environment on different scales?

Only the sequential development of a system of observations, of predicting the state of the environment, and of estimating the existing and predicted state of the environment, can answer the first question.

This report, and the symposium as a whole, are devoted to an attempt to answer the second question. The result of the answer to the second question is primary to estimating the state of the environment, and is, therefore, the scientific basis of monitoring; it is precisely this result that should be the basis for the optimization of the relationship between man and the environment, and the basis for monitoring and controlling the quality of the environment.

Most of the work in determining permissible loads up to this point has been limited primarily to the sanitary and hygienic approach (in developing permissible loads on man), and to the economic approach (the fisheries approach, for example). Secondary norms and standards too have been developed in which the possible consequences of effects on vegetation have been taken into consideration.

However, it is apparent that generalized approaches to the development of permissible effects on the population, of the assemblage of the system as a whole must be considered; individual sources must be taken into consideration, as must sources in this region, or system, as a whole, and the complexity of their effects, possible interactions. Moreover, it is desirable to find criteria to use to select, or define, the different variants of permissible effects (loads).

We shall, in making a complex analysis of all effective factors, estimate the final result (from the point of view of its acceptability) in terms of the magnitude of the damage. The latter can be broken down into ecological, economic, and moral (esthetic). Recreational damage can be included in the latter two kinds of damage. The kinds of damage indicated are interrelated to a considerable degree. Approaches to determining the extent of the damage will be described in what follows. It is desirable to adopt in advance certain strict limitations (for example, a total ban on any ecological or esthetic damage to unique objects, or, in general, any changes despite the fact that damage from a change at the particular point in time is not in evidence), as well as the unacceptability or unjustified (irreversible) ecological changes in most natural objects, and so forth.

Table 1 explains the kinds of damage that can be done, and establishes certain links between them.

Obviously, there is little chance of estimating esthetic damage at this time. There are certain approaches associated with an estimate of the time and money man spends to satisfy his need for esthetic values that can be taken to make such an estimate. Such expenditures are growing steadily, and will continue to grow in the future. Improvement, or preservation, of man's health will, at the same time, yield to evaluation.

The moral (esthetic) damage is determined to a considerable degree by an upset in man's habits, in his perceptions, emotions, and in the concepts he has formed during his upbringing, and as result of the education he receives during his lifetime.

We shall try to establish a link between the magniture of the source (emission, emissions) and the tude of the source (emission, emissions) and the scope of the ecological damage (factual $A_f^*$) within the framework of some limited area.

---

[*] First Deputy Chief of Main Administration of Hydrometeorological Service, Moscow

**Table 1**

| Ecological damage<br>A | Economic damage<br>B | Moral (esthetic) damage<br>C |
|---|---|---|
| I. Unjustified–irreversible consequences | Direct losses in the present; losses in the future | |
| II. Reversible consequences | | |
| a. deterioration in the condition of individuals, populations, ecosystem | Direct or potential losses in increase in biomass (harvest) | |
| b. response without deterioration in condition | | |
| –animals, plants | Possible migration with possible economic and *esthetic* losses | |
| –man | | Disruption of usual perceptions |
| | Recreational damage with possible economic and esthetic losses | |

For this we must know:

(1) the function of the sources (in space and time):

$$Q(\vec{R}, t^*) = \sum_{j} \sum_{l} \sum_{i} a_{jli}(\vec{R}, t^*) \qquad (1)$$

where

$a_{jli}$ is the different factors, ingredients, of $i^{th}$ orders within a homogeneous class;

j is the nature of the medium (air, soil, fresh, sea, or subsoil water, and so forth);

l is the nature of the heterogeneity (pollutant, radiation, heat);

(2) the geophysical laws of distribution (function F) in different media. The concentration (or value of some factor) $\sigma_{jli}$ in the different media is

$$\sigma_{jli}(\vec{R}, t) = F[Q(\vec{R}, t^*), v, K_{\vec{R}}, w_z, t\text{-}t^*] \qquad (2)$$

where

v is the rate of horizontal movement of an impurity (wind speed, flow, and so forth);

$K_{\vec{R}}$ is a coefficient characterizing turbulent diffusion;

w is a coefficient characterizing the vertical rate of transfer from one medium to another (sedimentation, precipitation, and so forth);

$t-t^*$ is the elapsed time between emission and time;

The possible transfer of part of a substance from one medium to another during distribution (migration) is

$$a_{j+1,li} = \mu\sigma_{jli} \qquad (3)$$

(3) the chemical (physical) laws for the conversion of a given ingredient (factor) into another

$$\sigma_{j,l+1,i+1} = \omega\sigma_{jli} \qquad (4)$$

where

$\sigma_{j,l+1,i+1}$ is the concentration (or value of the factor) of the ingredient at which the original ingredient, or factor, can be converted during distribution, or accumulation;

$\varpi$ is a coefficient characterizing the conversion;

(4) the distribution of organisms in space, $N_m(\vec{R},t)$;

m is their level; it is desirable to know $M_m$, the distribution of their weight relationship from the point of view of their ecological value, importance, determining their role in system stability. $N_m(\vec{R},t)$ takes into consideration the distribution of individual parts of the organism of a given specimen (population) in space (sometimes this is very important, as would be the case, for example, for plants, the roots of which are in the ground, while the above-ground organs are in the air).

The magnitude

$$\int_{\vec{R}} \int_{t} N_m(\vec{R}, t)\, d\vec{R}dt = U_m \qquad (5)$$

is an integral magnitude characterizing an m-population as a whole.

The importance of the factor $N_m(\vec{R},t)$ must be stressed, because the factual distribution or organisms in time and space determines to a considerable degree the ecological damage of an effect on the environment (because this distribution characterizes the possibility of contact between the organism and pollutants, or other factors).

For example, in estimating the factual situation in cities (determination of damage to the population) knowledge of the distribution of pollutants is not enough. This distribution must be compared with the distribution of the organisms in time and space in the particular city (including according to height);

(5) $\epsilon_{mj_l,i}$ is the measure of the biological effect (damage) for the given organism (level m) for the effect of a factor (ingredient) (l), entering the organism (or making contact) from the j medium. At the same time, it is very important to take into consideration the entrance of the particular ingredient fram all j media. The ingredient can affect different organs, so can differ in the effect produced as well.

The magnitude $\epsilon$, generally speaking, depends on the magnitude $\sigma$, on the rate of increase $d\sigma/dt$, and on its peak values, as well as on the duration of the effect.

In a more general form

$$\epsilon_{mjli} = \sigma_{jli}\, \delta_j(t)\, \tau_m \kappa_m \psi_m \qquad (6)$$

where

$\delta_j(t)$ is the percentage of the substance entering the organism in unit time (or of radiation irradiating the organism) from the medium including the geometric factor);

$\tau_m$ is the percentage of deposition, of immobilization in critical organs;

$\chi_m$ is the effect of the biological effect on any organ (on the organism as a whole) from the unit of immobilized substance (transmitted energy).

It is obvious that $\sigma_j\delta_j(t)\tau$ is the quantity of the substance in a critical organ (organs), or the quantity of radiation interacting with the organ as a function of time, and entering from medium j.

It should be pointed out that the magnitude $\epsilon$ depends on the condition of the organism. The latter is determined by a great many external effects, and undoubtedly by previous history (by the unhealthy, or weakened, condition of the organism) as well. This condition should be determined directly, or indirectly (by measuring the factors that influence the state of the physical and physicogeographical data, living conditions, and state of health), and should be introduced into the formula (the variable $\Psi$) used to characterize the condition of the organism, or population, or more correctly its deviation from the norm, attributable to the influence of external conditions, and including some distribution of the susceptibility of individual organisms in the given population.

Under normal, ordinary, conditions, the magnitude $\Psi_m$ can be taken as equal to unity (on the assumption of equal sensitivity of all organisms in the given population);

(6) $\nu_i$ is the effect of the interaction (the mutual effect factor (of two factors, i, k (ingredients), acting on the given organism;

The simultaneous effect must be estimated when several factors are acting (generally speaking, for all factors, including $\nu_m$).

The $\epsilon$ dependence is nonlinear, as a rule. The function can have dimensions (especially in the case of large doses, in the case of the destruction of the organism, of the onset in the organism of irreversible changes), possibly even in the case of very small doses (when the effect has a clearly expressed threshold) However, we are convinced that the effects of the action have no threshold (accumulation of data for small doses). This is important for summing effects from different ingredients;

(7) $\mathrm{Xi}_{m,n}$ is the result of the intensification (or weakening) of the effect on the given population, m, taking into consideration the effect of the different factors on several (n) populations simultaneously. The result of the complex effect may be written as $\xi_m$ where $\xi_m$ is a complex function of the values of $\xi_m$ and n.

Thus, the result of the summed biological effect $A_f$ is

$$A_f = \int_{\vec{R}} \int_{t} \sum_{m} \sum_{j} \sum_{l} \sum_{i} N_m(\vec{R}, t)\, M_m\, \epsilon_{mjli}(\nu_{i,k}\ \xi_{m,k})\, d\vec{R}\, dt \qquad (7)$$

It is desirable, in order to sum the biological effect of heterogeneous effects, to express the magnitude $\epsilon_{mjli}$ in units of permissible $\epsilon_{mjli}(p)$, or critical $\epsilon_{mjli}(cr)$ values. In other words, in terms of their relative values with respect to the permissible or critical effects for each type of effect, or ingredient (probably in $\epsilon_p$ units if $\epsilon \leqq \epsilon_p$ and $\epsilon_{cr}$, if $\epsilon_{cr} \geqq \epsilon > \epsilon_p$).

If the exact distribution, $N_m$, is unknown, the area of distribution, S(t), can be used conditionally. Incidentally, this factor can have a value, even if $N_m$ is known (it may be useful here to introduce the probability function for the potential "encounter" between the organism and the polluting impurity).

As already noted, what we have to know in order to resolve the question of the permissibility of the load on the system under examination, is the relative value of the biological (ecological) effect, not its absolute magnitude; that is, its relation to some permissible, or critical value. There are several approaches in the case of a comparison such as this: each magnitude $\epsilon_{mjli}$ can be compared to a previously processed value of $\epsilon_{mjli}(p)$ or $\epsilon_{mjli}(cr)$, which are magnitudes for use in establishing the values of the maximum permissible concentration (MPC) of ingredients for each organism and factor, or of the maximum permissible intake (MPI) (of the ingredient), if it is a question of the integral magnitude $\epsilon_{mjli}$ (t)dt in the case of a permissible load, or of values dangerous, or fatal, to the particular organism (population) in the case of a comparison with the critical magnitude, $\epsilon_{cr}$.

It is obvious that exceeding $\epsilon_{cr}$ for this population is absolutely unacceptable if we are talking about an individual population, but this assertion is not obvious for an individual population when $\epsilon_{m}(cr)$ is exceeded for the system as a whole.

This is followed by summing these relative values, using Eq. (7) (and taking $N_m(\vec{R})$ into consideration).

The result is the obtaining of $A_p$ and $A_{cr}$ (permissible and critical) values, developed for a specific, concrete, natural object, or system, that take into consideration all the peculiarities of the particular ecosystem (system), as well as the climatic and concrete (at a given time) natural conditions.

It would appear to be obvious that when we speak of the stability of biological systems (or of other systems), we have in mind criteria for establishing the critical effect (that can lead to upsetting stability, and then to the conversion, or destruction, of the system).

What we mean when we talk about developing an MPC, or MPI, for an individual organism, or population, is some definite threshold, or initial pathological (exceeding the limits of physiological norms) action, or effect, which usually is reversible, and which certainly does not upset the stability of the particular system.

Consequently, the formulation of an approach such as this for the entire system is extremely important. At this point it would appear that we can talk about MPC value for just a single, critical link. However, it is more important to formulate an integral approach, even if no single value reaches the magnitude of the MPC, but the sum of the effects exceeds some magnitude, $A_p$, for a system as a whole, acted on by a whole complex of effects simultaneously. This approach, in our opinion, should be basic to a comprehensive analysis of the environment.

The initial approach here can be the use of the assumption of the lack of a threshold in the accumulation of data (the occurrence of certain effects) and the determination of the magnitude $\zeta_p$, by summing the relative values, $\sum_j \sum_i \frac{\epsilon_{ji}}{\epsilon_{ji}(p)}$.

"A" will change by bounds, depending on whether the system is stable or will become unstable.

This formula also is suitable for determining damage when the effect is on a major system, when not only the ecological damage is critical (but perhaps not the major aspect, but so too is the possibility of upsetting the stability (or approaching this critical state) of a major physical (geophysical) system, when estimating damage that can result from possible climatic changes, for example.

Determination of the relative ecological damage, $A_f$, undoubtedly must be made successively, with possible approximation at first (selection of a limited number of factors for the effect of polluting ingredients, and assumptions, where possible). $\nu = 1$, $\Psi = 1$.

Systems analysis approaches are necessary when solving similar problems.

## The Scientific Justification for Monitoring

In order to estimate the effect on a population, community, or ecosystem by the scheme cited, we must have the values of all the variables contained in this formula. There is no question of the fact that the solution to the problem cannot be obtained in its entirety so we must proceed to simplify things by selecting the most important factors of the effect, and by searching for the links that are critical, or that govern the state of the systems.

Thus, regular observation must be made (variables $Q$ [illegible], $\dot{N}$, $\epsilon$, and others of the state of the medium, its future state must be predicted, we must be able to single out changes caused by anthropogenic action, and we must be able to properly estimate this action, by using the formulas cited above, for example. The values of $\nu$, $\chi$, and of some of the others, are found through research, and by setting up special experiments.

Table 2 shows the scheme for observation, for the preduction, and for the estimate, of the state of the environment (1). This system is, in essence, the scientific justification for monitoring. Let us explain the data in Table 2 in more detail.

Observations (I) of changes in the state of the environment include:

A.1.1. Observations of the state of the environment, and of changes in this state, characterized by geophysical data, are made by means of successive measurements of the corresponding parameters determining the instantaneous state of the environment. Such observations already are being made by a number of geophysical services.

A.1.2. Physicogeographic data, including data on the distribution of land and water, on the relief of the land surface, on natural resources (mineral, land, vegetable, water, fauna resources), population, urbanization, and others, also are most important information to have on the state of the natural environment.

A.2. Observations of the state of the environment (and of changes in this state) characterized by geochemical data, observations of matter and energy cycles in nature, observations of the composition of alien impurities in the biosphere (including radioactive ones), and observations of different specific types of pollution, including observation of noise, thermal pollution, and different types of radiation (the concepts of pollution in these latter cases are conditional, but are generally accepted).

**Table 2**

| | I | II | III |
|---|---|---|---|
| | Observation | Prediction | Estimate of existing and predicted states |
| A. State of environment | A. 1. State of the environment, characterized by geophysical and physicogeographic data $(N, v, K_{\overset{\rightarrow}{R}}, w_z)$ | | |
| | A. 2. State of the environment, characterized by geochemical data and data on the composition and nature of pollution $(\sigma_i, \omega, \mu)$ | | MPC[1] |
| | A.2.1. Local sources of pollution (Q) | | MPE[2] |
| B. Reaction of biota | B. 1. Responses and consequences: individuals and population $(\epsilon, \psi, \chi)$ | | MPC |
| | ecosystem (A)<br>biosphere as a whole (A) | | MPL[3] |

Criteria for permissibility of effect (in estimate):

1. MPC–maximum permissible concentration
2. MPE–maximum permissible emission
3. MPL–maximum permissible load

This section includes observations of the chemical composition (of natural and anthropogenic origin) of the atmosphere, precipitation, surface and subsurface waters, ocean and sea waters, soils, bottom deposits, vegetation, animals, and of the basic paths over which pollution is distributed.

Observations of local sources of pollution must be distinguished in this section (A.2.1.). These sources may be natural (primarily volcanic eruptions, as well as the spontaneous discharge of gases, oil, and so forth) and anthropogenic (industrial plants, agricultural sources such as stock-raising farms, fields after the application of chemicals to fight plant pests, and transportation, air, water, and land, and municipal sources).

B.1. Observations of the reactions of the biota, the living component of the biosphere, to changes in the state of the environment should include observations of the response (reversible changes) and consequences (irreversible changes). Observations are possible in terms of functional and structural biological criteria. The functional criteria can include the growth of the biomass per unit of time, as well as the rate of absorption of different substances by plants and animals. The structural criteria can include the number of species of plants and animals, as well as the total biomass.

Initial data on (1) the state of the environment (which can be obtained by implementing the first section of the observations), and on (2) plans for human activities, are needed to make the prediction (II). We can, when making predictions, assume constancy of human activity, or include in the calculation data on plans to increase the scope of economic activity (which can lead to an increase in environmental pollution), as well as measures for reducing environmental pollution. However, even a prediction based on the assumption of constancy of human activity can yield very valuable information on possible future changes in the biosphere.

Optimization of the relations between man and the environment will be achieved when man's activities, and his plans, are coordinated with recommendations flowing from the prediction and estimate of the state of the environment.

## Estimate of Possible Economic Damage, and Determination of Permissible Load

Economic damage, B, can be written in its most general form as $B = CA$, where C is a coefficient that depends on the nature of the system, and on concrete conditions. The magnitude C can simply signify the proportionality factor, or it can reflect dependency on certain economic factors, since practically all the ecological features of the given system are included in the value A.

Certain limiting conditions can be formulated (those situations in which, conditionally, $B \rightarrow \infty$, or in which such a value is impermissible);

(1) it is impermissible if A reaches $A_{cr}$, that is, values of A for which systems are converted into others, or are destroyed;

(2) it is impermissible for $A > A_p$ in unique objects, that is, for A to exceed some permissible value, $A_p$, which was determined above;

(3) moral damage is not now subject to estimation as a practical matter, so will not be considered here, or in what follows.

Let us designate as E conventional systems with a reserve (an "ecological" reserve for ecosystems)

$$E = A_{cr} - A_f \tag{8}$$

where

$A_f$ is the factual state of the system, the ecological load on the system; then

$$A_f = A_O + A_i \tag{9}$$

where

$A_o$ is the state of degradation of the particular system;

$A_i$ is the load attributable to the particular effect.

In our opinion, C should depend on E; that is, the damage increases with approach to the boundary of the ecological reserve, $C=E(A_{cr}—A_f)$; that is, a greater effort is required (including that of an economic nature) to return the system to a normal state from an unhealthy, or changed, one.

It is appropriate to introduce the concept of magnitude of damage per unit of production in order to make practical decisions

$$d_i = \frac{B_i}{D_i} = \frac{C(A_{i(cr)}-A_i-A_O)\,A_i}{D_i} \tag{10}$$

where

$D_i$ is a product of the $i^{th}$ type, manufactured in the particular region, leading to damage $B_i$; the subscript i should be replaced by the subscript $\Sigma$ when manufacturing the summed product to estimate the damage.

This formula is useful in making a decision as to the desirability (and feasibility) of manufacturing the particular product as a function of the permissible load, and of the relative ecological (or economic) damage.

The economic effectiveness of the measures taken to protect, or to restore, the quality of the environment may be estimated by comparing the assumed expenditures, G, with the economic damage, $B<B_{cr}$. These measures are effective if $r=G/B\leqq 1$. There is no doubt that the magnitude r depends on the time of implementation of these measures, because the magnitude A, and consequently B, is determined to a significant degree by time of effect.

## BIBLIOGRAPHY

1. Yu. A. Izrael, "A Global System of Observations, Prediction, and Estimating Changes in the State of the Natural Environment. Fundamentals of Monitoring." Report to the Intergovernmental Conference on Monitoring, Nairobi, February 1974.

# Defining "No Significant Deterioration" of Air Quality

Ralph A. Luken *

## I. INTRODUCTION

The Clean Air Act of 1970 requires the Environmental Protection Agency (EPA) to protect existing levels of air quality in non-degraded areas. Taken literally, protection means no change or deterioration in the existing levels of air quality. Since no deterioration would severely constrain economic growth in less developed areas of the United States, the court has interpreted protection to mean prevention of "no *significant* deterioration" (NSD) of air quality.[1] However, the court did not define *significant* in the NSD ruling and has left the task to EPA.

Since discussion in a short paper of all the complex issues associated with defining NSD is infeasible, I have chosen to concentrate on (but not limit myself to) developing an economic definition of NSD. The paper is divided into four major sections:

- Description of the Federal Law relevant to NSD
- Economic definition of NSD
- Critiqueing on the basis of the economic definition the operational definition of NSD preferred by
- The role of rigorous Environmental Impact Statements in determining the significance of environmental deterioration

## II. RELEVANT FEDERAL LAWS

### State Air Quality Implementation Plans

Under the Clean Air Act of 1970, EPA is charged with establishing two types of national ambient air quality standards: the national primary air quality standard, which is that level of air quality required to protect public health, and the national secondary ambient air quality standard, which is that level of air quality required to protect the public welfare from any known or anticipated adverse effects associated with the presence of such air pollution in the ambient air (Table 1).

EPA is also required to protect against significant deterioration of air quality even where that air quality exceeds the minimums prescribed in national standards. The statutory basis for a NSD policy in the 1970 Act is found in Section 101, which states that the purpose of the Act is "to protect and enhance the quality of the Nation's air resources." To allow deterioration of existing air quality in non-degraded areas is considered to be antithetical to achievement of the mandate to protect the quality of the Nation's air resources.

**Table 1**
**Primary and Secondary Air Quality Standards for $SO_2$ and TSP[a]**

($\mu g/m^3$)

| | $SO_2$ | | | TSP | | |
|---|---|---|---|---|---|---|
| | Annual. | 24 hrs. | 3 hrs. | Annual | 24 hrs. | 3 hrs. |
| Primary (Health) | 80 | 365 | ---- | 75 | 260 | -- |
| Secondary (Public Welfare) | -- | --- | 1300 | 60 | 150 | -- |

[a] There are also national ambient air quality standards for Carbon Monoxide, Photochemical Oxidants, Hydrocarbons, and Nitrogen Oxides.

The Act requires each State to submit to EPA a plan providing for the implementation, maintenance, and enforcement of such ambient air quality standards within each of its air quality control regions. As a result of the Sierra Club citizen suit, the plan must now provide for the prevention of significant deterioration of air quality in clean areas. Each plan must be approved by EPA before it becomes effective.

In summary, the Clean Air Act of 1970 establishes different standards for degraded and non-degraded areas. For degraded areas, the requirements are for achievement of primary and secondary air quality standards. For non-degraded areas, the requirement is for the protection of the existing level of air quality even though the quality level exceeds the standard.

## III. ECONOMIC DEFINITION OF NSD

Two heuristic economic definitions of NSD offer some insight into the problem of interpreting NSD. A general equilibrium definition would address the issue of whether national growth is over or under-constrained by a NSD policy. The definition in this case would be discussed in terms of a national linear programming model which seeks to minimize the costs of production subject to restraints specified for environmental standards and demands of the domestic market. The advantage of this approach is that it recognizes that development is of national as well as local concern and fully accounts for the interdependencies among several (if separate) locations.

---

*Chief, Water Programs Branch, Office of Planning and Evaluation, Environmental Protection Agency

[1] Sierra Club vs. Ruckelshaus, 344 F. Supplement 523 (D.D.C. 1972) affirmed, 4 ERC 1815 (D.C. Circular 1972).

A partial equilibrium definition would address the issue of whether a new source would be permitted in a specific location. The definition in this case would be discussed in terms of marginal damage and cost functions associated with the residuals generated by a plant. The advantage of this approach is that it focuses on the value of environmental damages as well as the cost of control. The partial equilibrium definition is used in this paper because of the advantages noted in the previous sentence as well as the fact that the damage and cost functions can be designed to account for some national as well as local costs and benefits.

Environmental economists have given some thought to the "efficient" level of discharge reduction. For example, they have examined the residuals discharged by industrial, municipal, agricultural, etc. activities and have concluded that the discharger must control his residual discharge up to the point where the additional costs of reduction begin to exceed the benefits from further curtailment of discharge.

The "efficient" level of discharge reduction for a degraded area [3] is diagrammed in the upper half of Figure 1.[4] The amount of residual discharge is recorded on the horizonal axis. In the absence of any control efforts by the polluter, the maximum level of discharge equals B. Also, the corresponding change in air quality is recorded on the horizontal axis; the highest level of environmental quality is 0. The cost of controlling residuals and the dollar value of damages associated with increasing levels of discharge are recorded on the vertical axis. Total damage cost (TDC) is shown for various levels of discharge along the curve OC. The curve has zero slope (essentially no marginal damages) from O to A on the assumption that there is no damage below the standard[5]; it increases at an increasing rate from A to C on the assumption that the damages (for example, health effects) increase with rising levels of discharge.[6] Total abatement cost (TAC) for a given level of product output is shown for various levels of discharge reduction along curve DB. The curve increases at an increasing rate because of the increasing marginal cost of controlling residuals. The cost of control is D if all residuals are controlled. Curve DHC is the total cost attributable to pollution and abatement, that is, it is the sum of TDC and TAC. It combines the cost of controlling residuals with the damage cost associated with the remaining residuals discharged. The efficient level of discharge reduction for a given level of output is discharge level E, where the sum is a minimum.

The efficient level of abatement may also be determined in the lower part of Figure 1 by comparing the marginal damage curve (MDC) and the marginal abatement curve (MAC). The efficient level of abatement is obtained when MDC = MAC as it is at E. To the right of E, the plant would be required by a society interested in efficient resource allocation to curtail its residual discharge either by regulatory actions or by the imposition of an effluent charge. To the left of E, the plant would be encouraged by a society interested in efficient resource allocation to discharge its residuals because the damage to the airshed is less than the cost of discharge reduction. However, the plant would be prevented from discharging its residuals in the range from G to E because the standards are based on protecting all populations against all known adverse health effects rather than a balancing of adverse health effects and the cost of discharge reduction.

The efficient level of abatement for a non-degraded area [7] is diagrammed in Figure 2. The units on the axes are the same. The shape of the TAC curve for a projected level of output is also the same. However, the shape of the TDC curve is fundamentally different as is the initial level of environmental quality. The TDC reflects the assumption on the part of many environmentalists that the initial deterioration or damage to a non-degraded area often results in a significant loss to aesthetic amenities and, if it is an ecologically fragile area, a significant loss in the regenerative capacity of the natural system. The shape of the curve suggests that the air quality standards are inadequate for protecting against all known long-term adverse effects. Thus the TDC curve in Figure 2 is a composite of damages because it reflects both aesthetic and ecological effects resulting from minute changes in ambient concentrations and the health and welfare effects normally associated with the TDC.

---

[3] The diagram follows A. M. Freeman, III, R. H. Haven, and A. V. Kneese, *The Economics of Environmental Policy* (New York: Wiley, 1973), p. 85.

[4] A deteriorated or degraded area is defined as an airshed where the level of ambient air quality violates the national standards (Table 1).

[5] The TDC is drawn in Figure 1 as a hockey stick (with a kink at A) to reflect the fact that all populations in all areas must be given the same level of protection against adverse effects.

[6] The TDC as depicted in Figure 1 is a short-run damage curve for one pollutant. A long-run damage curve which took into account cumulative effects, or a damage curve for several pollutants would probably have a gradual slope from O to A in Figure 1.

[7] A non-degraded area is defined as an airshed where the initial level of ambient air quality does not violate the national standards.

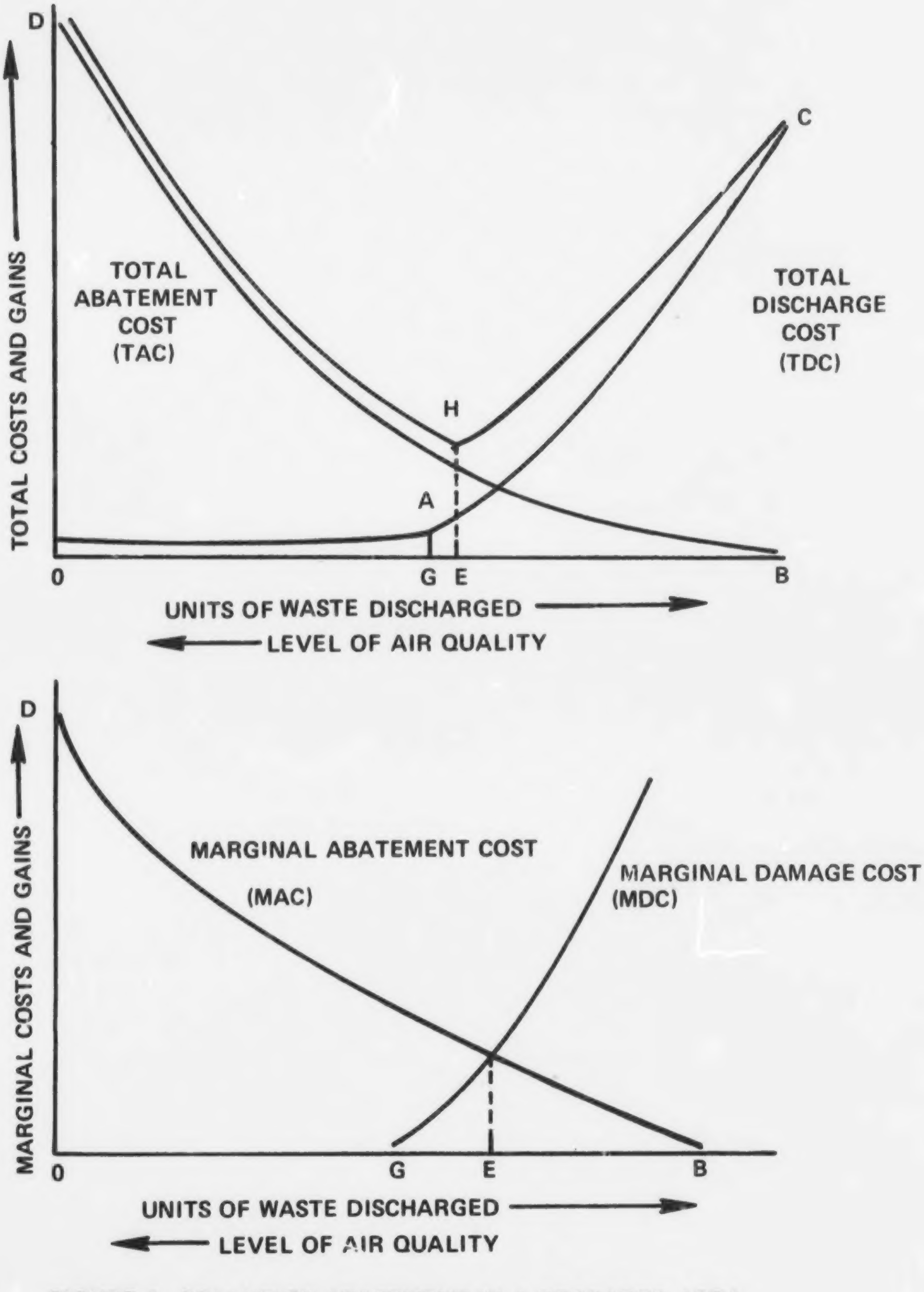


FIGURE 1. POLLUTION ABATEMENT IN A DEGRADED AREA

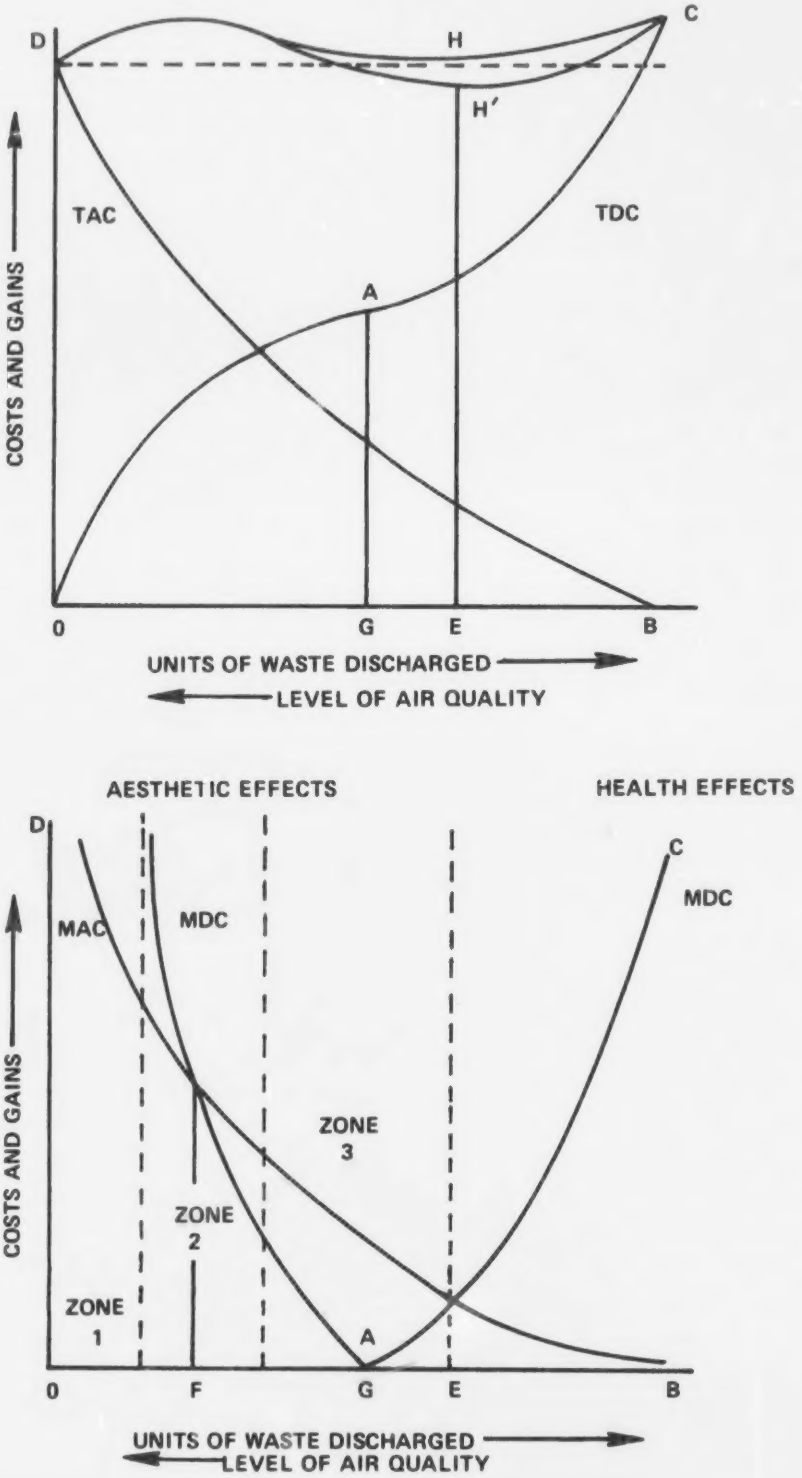


FIGURE 2. POLLUTION ABATEMENT IN A NON–DEGRADED AREA

The efficient level of abatement is determined in the lower part of Figure 2 by comparing the marginal damage curve (MDC) and the marginal abatement curve (MAC). The section of the MDC curve between I and A reflects the aesthetic and fragile ecological effects, while the section between A and C reflects the health and welfare effects. As long as the initial level of discharge is to the left of F, the efficient level of discharge is near zero. However, to the right of discharge level F, the efficient level of discharge would be in E in spite of the fact that damages to a non-degraded area would occur in the range from F to G and national standards would be violated in the range from G to E.

The MDC curve might shift from MDC to MDC′ in Figure 3a if it reflected national as well as local concern for non-degraded areas. Many urban residents in the U.S. have become increasingly interested in maintaining the environmental quality of non-degraded areas for their future use, for their descendants or just for the pleasure of knowing a non-degraded area is protected from environmental deterioration.[5] MDC′, which indicates that a higher value is placed on environmental damage, can be seen as reflecting both local and national concerns and thus is a more accurate indicator of societal value. Now as long as the initial level of discharge is at F′ or to the left of F′, the efficient level of abatement is to curtail residual discharge to near zero. However, to the right of discharge level F′ the efficient level of discharge reduction would be E because costs of control exceed demages in the range from F′ to E.

Similarly, the MAC curve would shift if there were opportunities foregone, such as electrical energy consumption by individuals outside the region. The argument is that cost curves ought to reflect both onsite and external costs rather than only on-site cost. If the users of electrical energy consumption outside the region were denied the benefits of lower cost electricity generated in the region, then their loss ought to be reflected in the MAC. MAC′ in Figure 3b, which indicates a higher cost of residual control, sums both on-site and external costs. In this case, there might not be a low level of discharge for which MDC would exceed MAC. The efficient level of abatement would be E′ rather than some low level of discharge for a non-degraded area as suggested in Figure 2.

The absence of certain types of information limits the operational usefulness of an economic definition. The most difficult problem is measuring the MDC function or, more simply, the value people place on clean air. Moreover, estimation of MDC is more than assigning values. For example, estimation of a MDC requires three sequential steps:

Step 1: construction of a transfer function which relates changes in residuals to changes in ambient conditions. Control of residuals can affect the concentration of such pollutants as particulate matter and sulfur oxides;

Step 2: estimation of the impact of changes in ambient concentrations on human health and welfare. Linking changes in ambient concentrations to human health and welfare is the most difficult step in estimating MDC;

Step 3: Identification of the economic value of induced changes in human health and welfare.

On first thought one would think about asking people how much they value changes in clean air to obtain information about Step 3. Since people are likely to understate their preference for jointly-shared social programs, several indirect techniques have been devised to overcome this limitation. These techniques measure losses in wages, residential property values, materials, and vegetation. While these techniques have provided some insight into the benefits of pollution control (damages avoided), they usually do not generate sufficient monetary information on which to base a MDC function.

Measuring the cost of control is sometimes as difficult as measuring the damages avoided. If several plants are required to control their residuals simultaneously, then the change in demand for pollution control equipment may alter the price of this equipment for all plants. If a plant controls its residuals by transferring them from water to the air, then the control cost must also include the damage imposed on the other medium. If there were an opportunity cost associated with a high level of abatement, that is, there are benefits foregone by inhabitants in other regions because the final output is restricted, then the control cost must reflect the value of these foregone benefits.

Also, there are the usual problems of separating pollution control costs from process change costs and determining if the value of the controlled pollutants as a secondary product will offset the cost of control. Information about all these conditions is important for a correct assessment of costs.

Another difficulty is estimating the initial level of environmental quality. In many non-degraded areas there are few if any monitoring stations for TSP and $SO_2$. Vast numbers of measurements would be required to establish the baseline level of environmental quality. Data variability is still another problem. Normal variations in pollution concentration in clean areas, especially TSP, exhibit considerable variability.

---

[5] There is considerable literature on the non-user values for environmental quality (Krutilla, 1967). Weisbord (1964), for example, has made a strong case for inclusion of option value in the total value of a service. Option value is the value individuals place on a service when they are uncertain about their future need for it.

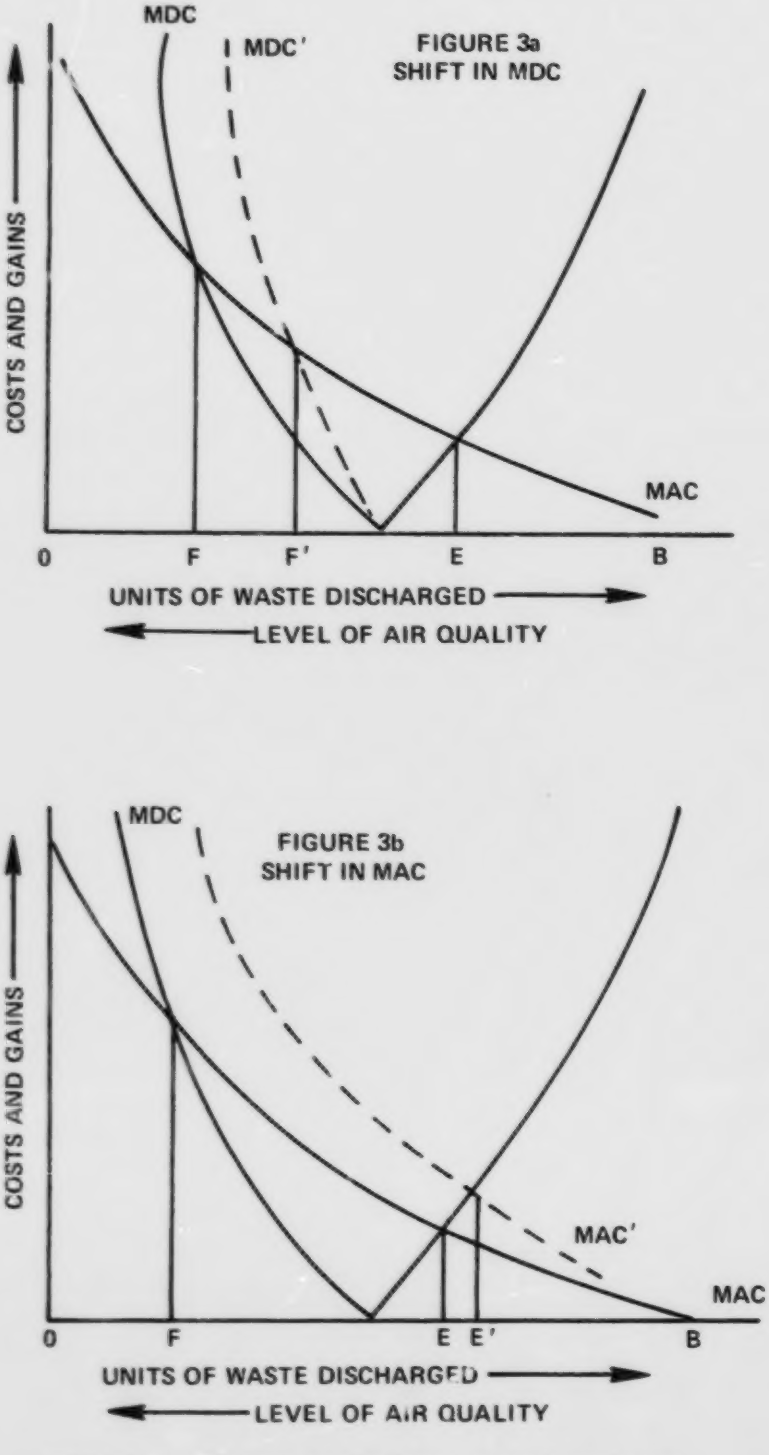


FIGURE 3. EFFECT OF SOCIETAL EVALUATION ON MARGINAL COSTS

For example, the 1968 TSP maximum concentration at the Grand Canyon for particulates was 126 $\mu g/m^3$, and the annual average was 31 $\mu g/m^3$. In 1969, the maximum concentration was 32 $\mu g/m^3$, and the annual average was 12 $\mu g/m^3$. These differences are caused by random variations due primarily to normal meteorological factors.

Even if essential data are unavailable, framing the issue of NSD as an economic problem is useful because it indicates the conditions under which varying degrees of deterioration of environmental quality are efficient and are inefficient. Formulation of the issue as depicted in Figure 2 suggests that permissible changes in air quality in non-degraded areas ought to vary with the original level of ambient quality and the shape of the costs and damages curves. Essentially, permissible changes will fall into three categories:

- no change from current ambient concentration would be permitted under conditions where there is exceptionable ambient quality and MDC is significantly greater than MAC (Zone 1 in Figure 2).[9]
- some change from current ambient concentrations would be permitted under conditions where there is high ambient quality, but uncertainty about whether MDC is greater or less than MAC (Zone 2 in Figure 2). As indicated above, there may be considerable uncertainty about the shape and thus intersection of the MDC and MAC curves.
- as much change from current ambient concentration as necessary to accommodate new development under conditions where there is good ambient quality and relative certainty that MAC is greater than MDC (Zone 3 in Figure 2). The permissible increment would vary depending upon whether national standards dictated the upper limit (point H, Figure 2) or whether an efficient level of abatement dictates the upper limit (point E, Figure 2).

In summary, a definition of NSD consistent with efficient resource allocation must meet two conditions. First, the definition must reflect both the costs and damages associated with a change in residuals discharged. A definition based on either costs or damages alone would not result (unless by chance) in an efficient level of residuals discharged. A definition of NSD based only on a change in ambient quality (damages avoided) is inadequate because it ignores the costs of achieving that change. A definition of NSD based only on a change in emission rates from a plant is inadequate because it ignores the change in damages associated with that change. Both changes in ambient conditions (damages) and emission rates (costs) are central to establishing a definition of NSD consistent with economic theory.

Second, the definition must reflect the unique characteristics of the situation. Both the original level of ambient quality and the shapes of the costs and damages curves will probably vary from region to region. The permissible, i.e. non-significant, change will vary within a non-degraded area and among non-degraded areas.

## IV. OPERATIONAL DEFINITION OF NSD

The operational definition of NSD of air quality preferred by EPA is partially but not entirely consistent with the economic definition of NSD. The definition provides that:

- States divide their territory into three zones;
- the Federal Government establish for each zone deterioration changes which cannot be exceeded.

### Zone Designation

A Zone I designation would be areas where little if any change from current air quality patterns is permissible. States would probably use Zone I designation where:

- Development of industry within the zone is inconsistent with desirable uses for the area.
- The area should be protected from degradation due to pollution sources outside the area.
- Development of polluting sources would significantly degrade an area with very clear air.

A Zone II designation would be relatively undeveloped areas where some change is permissible. A Zone II designation would be a recognition that the point of transformation between Zone I and Zone III in a non-degraded area is often difficult to distinguish because of the uncertainty about the cost and damage functions and the absence of physical data. This designation would prevent the arbitrary designation of an area as either Zone I or III until sufficient data were available to make a decision.

A Zone II designation would be areas where major industrial growth is desired and where increases in concentrations up to the national standards would be insignificant. (Again, note that the national standard as a limitation is inconsistent with the economic definition, which sets the point where MAC = MDC as a limitation.) States probably would use Zone III designations where:

- an existing urbanized region has air quality which does not violate secondary standards.

---

[9] This definition appears to be more restrictive than the statements of the Sierra Club on the issue of NSD. The Sierra Club recognized that NSD does not prevent all increases in pollution. "If the best available technological developments are utilized and if numerous pollution-producing sources are not concentrated in one place, most industry can enter clean areas without causing significant deterioration." *Federal Register* July 16, 1973, p. 18987.

- counties are urbanizing.
- areas are designated for concentrated development.

## Zone Changes

The Federal Government has indicated permissible air quality changes for Zone I and II for $SO_2$ and TSP (Table 2). (A permissible change for Zone III is not necessary because it is the difference between initial ambient quality and secondary standards.) The permissible changes for Zone II were calculated backwards from the impacts on the ambient environment in an open terrain of an 850 MW power plant and a 200,000 bbls/day petroleum refinery with the application of best practicable control technology. (See Appendix A for more detail.) The permissible changes for Zone I were very small numbers, arbitrarily chosen, which would allow little or no change from the initial ambient quality.

**Table 2**
**Permissible Increments for Zones I and II**

| | $SO_2$ | | | TSP | |
|---|---|---|---|---|---|
| | ($\mu g/m^3$) Annual | 24 hrs. | 3 hrs. | ($\mu g/m^3$) Annual | 24 hrs. |
| Zone I | 2 | 5 | 25 | 5 | 15 |
| Zone II | 15 | 100 | 300 | 10 | 30 |

Source: Harbridge House, "The Impact of Proposed Non-degradation Regulations on Economic Growth." EPA, BOA #68-01-1561, November, 1973, Volumes I-III. Also, U.S. Federal Register (July 16, 1973), pp. 18990–18993.

The permissible changes for Zone I and II are not absolute limits on the potential changes in ambient concentrations, in that State government can alter the final designation of a given area if the resulting air quality does not correspond to significant degradation.

The area classification definition is partially but not entirely consistent with the economic definition for the following reasons. First, it incorporates the idea that the permissible deterioration in air quality ought to vary within a non-degraded area and among non-degraded areas. It allows for variability within a non-degraded area with the three zone classification which takes into account the initial ambient concentration and the physical, biological, and economic development status of an area.

It allows for variability among non-degraded areas by permitting States to modify the level of deterioration given adequate consideration of costs and benefits. Second, the permissible changes established by EPA are based on the cost of control. The Federal definition is based upon technology which has already been installed by industrial establishment without undue financial hardships. In addition, an economic analysis of the impact of the Zone II requirements suggested that industrial growth would not be unduly restricted in upper New York State, Dallas/Fort Worth area, and the Four Corners region.[10]

The definition is inconsistent with the economic definition in that the operational definition was not based on a comparison of the benefits and costs associated with the zonal limitations. The zonal limitations incorporated only the first step in benefit analysis, i.e., relating residuals discharged to changes in ambient concentrations. It did not incorporate the effect of the changes in ambient concentrations on human health and welfare and the economic value of the changes in human health and welfare. Thus, there is no information about the physical and monetary damages associated with the permissable changes for Zone I and II. Whether they are of greater or less damages than the associated costs of control is still an unanswered question.

In summary, the proposed operational definition is only partially consistent with the economic definition of NSD. The operational definition accounts for the need for variation in permissible air quality changes in non-degraded areas and is based on the costs of controlling residuals. However, the permitted changes in air quality for each zonal type are not based on the damages or the absence of damages associated with the permissible changes; this will restrict the ability of the States to balance costs and benefits to some extent in their zonal designation.

## V. THE IMPORTANCE OF ENVIRONMENTAL IMPACT STATEMENTS

As a practicable matter, estimation of a national MDC function associated with the permissable changes is an impossible empirical task. The data are not uniformly and consistently available. However, Environmental Impact Statements (EIS), which are micro-evaluations of environmental consequences of proposed actions, are a potential source of data about local damages. They potentially offer data for estimating the local MDC associated with the residuals generated by a new industrial source.

The EIS, which was introduced into the Federal decision-making process by the National Environmental Policy Act of 1969, must cover five points:[11]

Part 1—the environmental impact of the proposed action,

Part 2—any adverse environmental effects which cannot be avoided should the proposal be implemented,

---

[10] Harbridge House, "The Impact of Proposed Non-Degradation Regulations on Economic Growth." EPA BOA #68-01-1561.

[11] The following material is based on a paper prepared by Neil Orloff, "The Environmental Impact Statement Process," Director, Regional Liaison Staff, EPA, Washington, D.C. February 1973.

Part 3—alternatives to the proposed action and the results of not accomplishing the proposed action,

Part 4—the relationship between local short-term uses of man's environment and the maintenance and enhancement of long-term productivity, and

Part 5—any irreversible and irretrievable commitments of resources which would be involved in the proposed action should it be implemented.

The proposed actions under Part I are limited primarily to three types of action—licensing activities, funding activities, and activities directly and wholly undertaken by the Federal Government. More recently the courts have ruled that reconstruction of a segment of highway was a Federal action because, even though that segment was not Federally-funded, the overall project was Federally-funded.

The environmental effects under Part 2 have been defined very comprehensively in the Federal guidelines. The environment is not to be considered just in terms of air or water pollution or physical changes in the land. The environment also includes the social environment. The environmental effects must take into account the adverse effects in *excess* of those created by existing uses in an area. Thus the effects in a non-degraded area could be recorded as being very different than the effects of a similar ambient change in a degraded area. Also, environmental effects must take into account the cumulative harm to an environment.

Consideration of alternatives under Part 3, as this process has come to be defined by Federal guidelines, could contribute significantly to estimating a MDC:

> "Agencies must indicate in their procedures that all reasonable *alternatives* and their environmental impacts are to be discussed, including those not within the authority of the agency. Examples of specific types of alternatives that should be considered in connection with specific kinds of actions should be given where possible. Such examples should include, where relevant:
>
> (1) the alternative of taking no action;
>
> (2) alternatives requiring actions of a significantly different nature which would provide similar benefits with different environmental impacts (e.g. : fossil fuel vs. a nuclear power plant);
>
> (3) alternatives related to different designs or details of the proposed action, which would present different environmental impacts (e.g. pollution control equipment on a nuclear plant).
>
> In each case, the analysis of alternatives should be sufficiently detailed and rigorous to permit independent and comparative evaluation of the benefits, costs and environmental risks of the proposed action and each alternative." [12]

Thus, the process of analyzing alternatives has the potential for generating not only physical data about a project, but also data about the value associated with a change in the environment.

However, the greatest potential for providing information about how the public values environmental damages reported in an EIS is in the public comment period, assuming that there is well-prepared EIS. For the first time in many instances, concerned citizens will have the technical information on the environmental consequences from the introduction of a new source of residuals and an opportunity to state formally their opinion. Draft EIS's must be available to interested public and private groups for comments and any substantive comments must be included in the final EIS. This process allows them to record their valuation of the environmental impacts.

Since public and private groups in all parts of the U.S. can comment on EIS's, the comment process has the potential for soliciting local, State, regional and national concerns. If there are numerous expressions of non-local concern, they would provide at a minimum an indication that significant values are being lost and at a maximum a basis for a shift in the MDC depicted in Figure 3a. The new MDC would be a more accurate indicator of value because it would be a summation of the value of all interested parties for a jointly shared good.

This type of information for estimating MDC will be generated for more projects in future years because there is a growing trend in the States to adopt requirements for EIS's like those of the Federal Government. Fifteen States and Puerto Rico now require impact statements for a wide range of activities (sometimes including local and private projects) significantly affecting the quality of the environment, and several others apply the process to limited classes of projects. At least twenty other States have such requirements under consideration.[13]

In concluding this section, I do not want to leave an overly optimistic impression about how much data EIS's will provide for estimating MDC. Estimation of a total damage cost curve, from which the MDC is derived, requires physical and economic data over a range of abatement levels. Physical and economic data over such a range are very difficult if not impossible to obtain. However, EIS's do have the potential for generating a few physical data points and a proxy for the value of the environmental damages. As limited as this information might be, it would be more information than is usually available for determining whether a new source should be permitted in an area.

---

[12] U.S. Council on Environmental Quality, "Guidelines on the Preparation of Environmental Impact Statements," 40 *CFR* 15.8A4.

[13] A more extensive description of State programs is available in Thaddeus Trzyna, "Environmental Impact Requirements in the States," EPA Contract #68-01-1818 (1974).

## VI. CONCLUSION

The proposed economic definition of "no significant deterioration" suggests that the permissible, that is non-significant, changes in ambient air quality vary within and among non-degraded areas. The size of the permissible change would depend upon the initial level of ambient quality and the costs and damages functions.

The operational definition of NSD preferred by EPA would require States to classify their territory into three zones and the Federal Government to specify permissible changes in each zone. The area classification definition is partially consistent with the economic definition. It is consistent because it incorporates the idea that the level of permissible changes ought to vary within and among different areas and it is based on the cost of control. It is not entirely consistent because it is based on a limited comparison of the benefits and costs associated with the zonal limitations.

As a practical matter, estimation of a national damage function associated with the permissible changes is an impossible empirical task. One viable option, however, is to estimate the local MDC resulting from the residuals generated by a new industrial source. Environmental Impact Statements, which are micro-evaluations of environmental consequences, are a potential source of data for this task.

## BIBLIOGRAPHY

Anderson, F., *NEPA in the Courts* (Baltimore: The Johns Hopkins Press, 1973).

Freeman, A. M. III, Haveman, R. H. and Kneese, A. V., *The Economics of Environmental Policy* (New York: Wiley, 1973).

Harbridge House, "The Impact of Proposed Non-Degradation Regulations on Economic Growth," EPA, BOA #68-01-1561, November 1973, Volumes I–III.

Krutilla, J., "Conservation Reconsidered," *American Economic Review*, 57 (December 1967).

Orloff, N., "The Environmental Impact Statement Process," speech, February 1973.

Sierra Club vs. Ruckelshaus, 344 F. Supplement 423 (D.D.C. 1972) affirmed, 4 ERC 1815 (D.C. Circular 1972).

Trzyna, T., "Environmental Impact Requirements in the States," EPA Contract #68-01-1818 (1974).

U.S. Department of the Interior, "Navajo Project Environmental Statement," September 1971, draft.

____________________ "Navajo Project Environmental Statement," February 1972, final, Appendix I (Review Comments).

____________________ Statement on Anti-Degradation, September 1968.

U.S. Federal Register (July 16, 1973), p. 18987.

Weisbrod, B., "Collective Consumption Services of Individual Consumption Goods," *Quarterly Journal of Economics*, 78 (August 1964).

## APPENDIX A
## TECHNICAL BASIS FOR ZONE II PERMISSIBLE CHANGES IN AMBIENT AIR QUALITY

Table A-1 shows the impact of five point sources on four Air Quality Control Regions. Concentrations were computed after the application of Best Practicable Control Technology. High concentrations caused by the power plant are due to several assumptions. These assumptions called for an 850

**Table A–1**
**Predicted Impact ( g/m$^3$) of Point Socrces in Four ACQR's***

| | *RURAL* | | | | | | | | *URBAN* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Open Terrain* | | | | *Valley Terrain* | | | | *Open Terrain* | | | | *Valley Terrain* | | | |
| | *SO* | | *TSP* | | *SO* | | *TSP* | | *SO* | | *TSP* | | *SO* | | *TSP* | |
| Source Type | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual | 24-hr. | Annual |
| By-Product Coke Oven | 1.6 | 0.3 | 0.7 | 0.6 | 2.4 | 0.3 | 2.1 | 0.6 | 0.6 | 0.2 | 0.8 | 0.6 | 2.4 | 0.2 | 2.1 | 0.7 |
| Cement Plant | .. | .. | 0.7 | 0.1 | .. | .. | 0.5 | 0.1 | .. | .. | 0.7 | 0.1 | .. | .. | 0.4 | 0.1 |
| Petroleum Refinery | 110 | 23 | 15 | .. | 122 | 26 | 25 | .. | 100 | 16 | 20 | .. | 122 | 26 | 50 | .. |
| Power Plant | 95** | 13 | 0.6 | .. | 437** | 16 | 0.8 | .. | 89** | 8.7 | 0.6 | .. | 447** | 9.9 | 0.8 | .. |
| Municipal Incinerator | .. | .. | 0.1 | .. | .. | .. | 0.1 | .. | .. | .. | 0.1 | .. | 0.1 | .. | 0.1 | .. |

*AQCR: Air Quality Control Region
**High concentrations due to short stack (122 meters) and large size (850MW)
NOTE: The figures in this table represent those maximum ground level concentrations estimated to accompany the *Best Level of Technology*.

MW plant, burning 2% sulfur coal with a stack of 122 meters. If the stack height had been greater, the contribution would have been less.

Table A-2 shows $SO_2$ and TSP concentrations resulting from 14 power plants in the midwest United States. Twenty-four hour results are based on observed stability wind readings obtained from meteorological observations at nearby airports for a period of 1 year. Existing output of $SO_2$ and TSP were modified to reflect compliance with New Source Performance Standards (NSPS).

Table A-3 shows 1-hour $SO_2$ concentrations around five coal-fired power plants. Column 4 shows measured $SO_2$ and column 5 modified $SO_2$ concentrations which are likely if the plants were in compliance with NSPS.

### Table A–2
### Maximum 24-Hour Concentrations for Selected Power Plants Using Observed Stability Wind Roses*

| Plant Number | Capacity (MW) | Volume ($M^3$/sec.) | Stack Gas Temp. (°F) | Stack Height (m) | 24-hr. SO ($\mu g/m^3$) | 3-hr. SO ($\mu g/m^3$) | 24-hr. TSP ($\mu g/m^3$) |
|---|---|---|---|---|---|---|---|
| 1 | 444 | 177 | 430 | 91 | 52.8 | 160.03 | 7.5 |
| 2 | 414 | 182 | 410 | 76 | 52 | 227.36 | 7.8 |
| 3 | 53 | 25 | 500 | 31 | 24.4 | 75.4 | 10.2 |
| 4 | 75 | 60 | 430 | 60 | 9.5 | 52.78 | 2.4 |
| 5 | 908 | 130 | 400 | 91 | 106.4 | 513.3 | 13.7 |
| 6 | 133 | 88 | 440 | 44 | 54.3 | 309.72 | 5.6 |
| 7 | 221 | 100 | 436 | 92 | 18.5 | 88.16 | 3.5 |
| 8 | 450 | 456 | 415 | 168 | 21 | 48.72 | 1.9 |
| 9 | 531 | 1,218 | 361 | 152 | 36.1 | 27.0 | |
| 10 | 100 | 51 | 405 | 69 | 17.3 | 71.0 | 5.1 |
| 11 | 396 | 210 | 460 | 76 | 79.3 | 175.16 | 4.9 |
| 12 | 308 | 70 | 450 | 65 | 57 | 356.7 | 8.1 |
| 13 | 59 | 200 | 450 | 83 | 9.5 | 55.1 | 0.59 |
| 14 | 1,304 | 440 | 422 | 208 | 56 | 137 | 1.2 |

*Existing coal burning plants, maximum estimated ground level concentrations assumed to meet NSPS by burning 0.7% sulfur coal.

### Table A–3
### Impact of Selected Power Plants on 1-Hour Air Quality*

| *Source Description* | *Fuel* | | *Measured 1-hr. SO* | *Modified 3-hr. SO* | *Distance to Max Concentration (km)* |
|---|---|---|---|---|---|
| | *Present* | *Assumed* | | | |
| 1000 MW Plant<br>Flat Terrain<br>2–450 ft. Stacks | 1.5%S<br>Coal | 0.7%S<br>Coal | 1048 $\mu g/m^3$ | 273 $\mu g/m^3$ | 5–6 |
| 1500 MW Plant<br>2–265 ft. Stacks<br>10–250 ft. Stacks | 3%S<br>Coal | 0.7%S<br>Coal | 2358 $\mu g/m^3$ | 319 $\mu g/m^3$ | 4 |
| 120 MW Plant<br>2–265 Ft. Stacks | 1.3%S<br>Coal | 0.7%S<br>Coal | 733 $\mu g/m^3$ | 273 $\mu g/m^3$ | 3 |
| 455 MW Plant<br>Rolling Terrain<br>3–400 ft. Stacks | 2.2%S<br>Coal | 0.7%S<br>Coal | 825 $\mu g/m^3$ | 150.8 $\mu g/m^3$ | 3 |
| 1700 MW Plant<br>Flat Terrain<br>2–600 ft. Stacks<br>1–800 ft. Stack | 3.8%S<br>Coal | 0.7%S<br>Coal | 2600 $\mu g/m^3$ | 304 $\mu g/m^3$ | 18 |

Maximum estimated average ground level concentrations

# Determining Acceptable Levels of Health and Environmental Damages

Fred H. Abel *

## I. INTRODUCTION

This conference is concerned with the process of developing comprehensive environmental quality management plans. These plans include a determination of the maximum pollutant environmental loading (MPEL), i.e., a determination of the maximum emissions allowed in any part of the region at any point in time. The determination of the MPEL, in turn, depends upon the level and kind of economic activity allowed, the cost society is willing to pay to obtain the desired level of environmental quality, the level of damages from pollutants society will accept, the distribution of costs (who pays costs) and the distribution of benefits (who receives benefits).

Societies must realize that the determination of MPEL does depend upon these factors and so these factors should be explicitly considered in the decision processes. In deciding to control pollution, societies have decided to trade off some economic growth, less expensive products, or better quality products for a better quality environment. This process of considering and then trading off the advantages of additional pollution control against the cost of additional control is called cost-benefit or trade-off analysis.

Whenever a decision is made, there is, either implicitly or explicitly, consideration given to the adverse effects (costs) and desirable effects (benefits) that will result from the decision. Our objective as scientists is to make these trade-off considerations as explicit and as full of knowledge as possible.

The presentation will proceed as follows. The concept of damage functions is developed with attention being given to the kinds of environmental pollution costs and health damages. Problems of quantifying damage functions and in turn valuing them in economic terms are discussed. The concepts of social costs of control and cost functions are developed. The principle of comparing costs to benefits is demonstrated in terms of economically valued functions. This process of balancing costs with benefits is extended to include complications arising when many costs and benefits cannot be valued in economic terms. Finally, the status of cost-benefits as a criterion for determining acceptable levels of pollution in the United States is presented.

## II. DAMAGE FUNCTIONS

There are many adverse effects of emitting pollution into the environment. Most societies are concorned with the health and environmental damages from pollution.

The number and kind of effects society suffers from pollution are a function of the level of pollution and the exposure of people to it. Curve A in Figure 1 shows the number of effects increases as the level of pollution increases. The effect function measures the rate of incidence of a specific effect, such as a premature death, an asthma attack, wildlife destroyed, crops destroyed, etc. Similar functions would be determined for each kind of effect.

Curve A in Figure 1 is a hypothetical curve: the actual curve could be quite different for each pollutant and each kind of effect. If a pollutant causes few effects below a given level (threshold) but many effects above the level, then the damage function would look more like Curve B in Figure 1. The acceptable level of pollution depends upon the exact slope of the damage function and in turn, on the location of point P1, if there is a threshold. In practice, we may have a function like A but, because only the part of the function above point P3 has been determined due to the difficulty of monitoring and determining effects at low levels of pollution, point P3 may be misinterpreted as a threshold.

The functions A and B in Figure 1 imply a very precise relationship between the level of exposure and the number of effects. Actually, a statistical procedure is needed to estimate these curves. Thus, a point on a curve, like point Z, is only the average or the expected number of effects. There is a frequency distribution of number of effects from exposure level P2 around the point Z as indicated by the curve H. There are two reasons for the frequency distribution

---

*Economic Analysis Branch, Washington Environmental Research Center, U.S. Environmental Protection Agency

Preceding page blank

H. One is the natural variation in individuals' or environmental response to pollution exposure P2. Second, there are errors in the data used to quantify the relationship. The variance becomes relatively large at low levels of exposure, primarily because of the data problem.

If the variance of the function A is known, then confidence limits can be estimated and confidence bounds like A' and A" can be displayed as in Figure 1. The existence of variance around the damage function makes the process of determining the acceptable level more complex and less certain.

The damage function is based on exposure and not on emissions. This has a significant impact on regional environmental management because exposure can be reduced in several ways. One way is to reduce the level of pollutants emitted into the environment. A second way is to shift the location of either the receptors (the people) or the polluters, such that exposure is reduced. A third way is to reduce exposure by removing the pollutants from the environment at the receptor, i.e., air conditioning buildings to avoid air pollutants, or pretreating water to remove pollutants.

## III. ENVIRONMENTAL POLLUTION COSTS

Pollution imposes burdens (pollution costs) upon society by causing deterioration in materials and living things (damages), by causing concern and dissatisfaction with the quality of the environment (psychic costs), and by causing extra expense (avoidance costs) to reduce the potential damage and psychic costs.

### Damage Costs

Included in this category are those effects on physical things, including human health, that result from exposure to pollutants. Effects on materials, buildings, vegetation and direct expenses involved with adverse health effects are examples. These damages result in decreased value of an item, decreased useful life of a structure or item and direct out-of-pocket costs associated with the damage (i.e., doctor or medicine costs). Health damages will be discussed in some detail below.

In estimating *damage* from pollution, dose-response relationships that show how much more

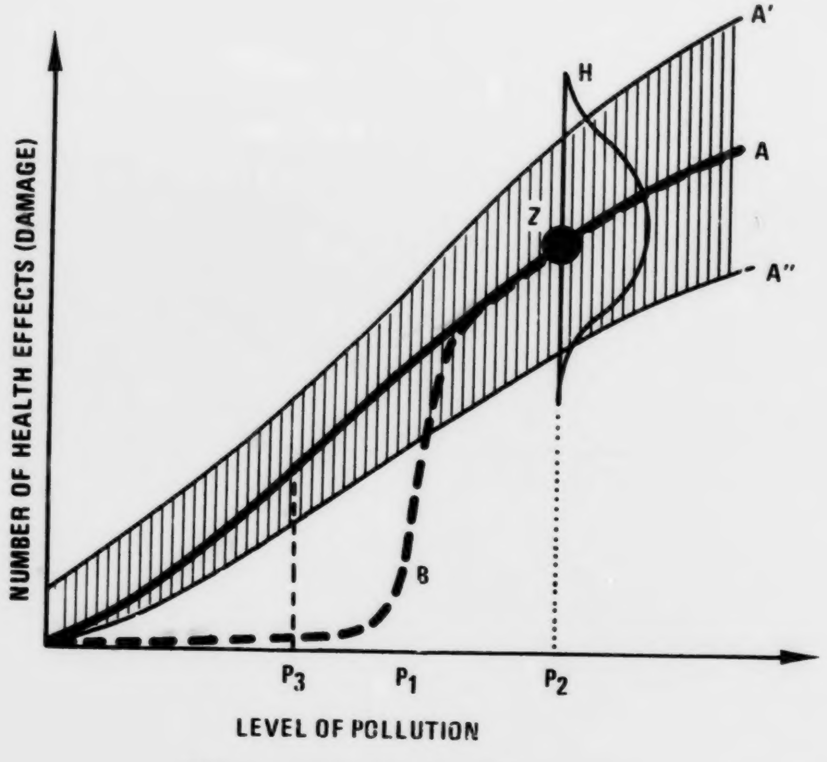


FIGURE 1. DAMAGE FUNCTION

quickly things deteriorate and become nonfunctional as a result of being exposed to different levels of pollution are needed. Dose-response relationships will allow estimates of the extra costs to industry and to society for replacing equipment and buildings more frequently. In the United States, estimates have been made of the level of materials damage from several air pollutants. These estimates are shown in Table 1. (Blank cells indicate that estimates are not available.) This was done by determining the levels of exposures and quantities of materials exposed under the current situation and estimating how much less deterioration there would be if pollution levels were reduced to meet the standards. The materials damage estimates are for unpainted metals, paint, rubber, and textiles. It was determined that the economic values of damages to many other materials such as concrete, glass, stone, and leather are very small. Estimates of damage to vegetation were determined in a similar way. That is, estimates were obtained of: (a) how the yield of agricultural crops or the growth of horticultural plants are related to levels of pollution, (b) the total acreages of these commodities exposed to pollution and (c) the value of the damage. The damage value was assumed to be the loss in revenue to producers because of reduced yields.

## Psychic Costs

This category includes the value of dissatisfaction and reduced quality of life that results because pollution exists. It includes the value of the dissatisfaction that results because (a) pollution interferes with enjoying the natural environment, and with visibiilty of scenic vistas, (b) populations of wildlife and natural vegetation have been reduced, (c) pollution causes discolored vegetation, dirty buildings and littered waterways and (d) people know the environment is being violated by indiscriminant dumping of wastes. It includes the value people associate with maintaining as many natural or high quality areas as possible because this gives them more choices for recreation. It includes the value people place on leaving future generations a better quality environment, and it includes the value of fear that people have for the welfare of the environment and the welfare of mankind. People are willing to live with some risk. At some point, they will be willing to give up some goods and services to reduce risk. Determining this level of acceptable risk is a major need in environmental decision making.

Estimating psychic costs is very difficult. It is almost impossible to obtain a dose-response relationship because the response is primarily in the state of mind of a person. A dose-response relationship is obtained either by observing the behavior of people as changes in environmental loadings occur, or getting them to express their changed state of mind in some terms that can be quantified.

For example, if particulate matter is emitted into the air, it can cause soiling of buildings and property; and if people find this reduction in aesthetic quality of the buildings objectionable, they would spend additional energy, time and resources in cleaning up the properties. The level of effort expended would be a measure of the damage or social costs from particulate matter emissions. Another method of estimating psychic costs associated with aesthetics in urban areas is concerned with the value of property or the level of rent. In either case, the value of the property or the value of the rent will be reduced if people perceive an aesthetically less pleasing environment. The aesthetic values in Table 1 were derived in this manner.

The extra cost of traveling to alternative recreation sites would be an estimate of the value of the psychic costs of pollution. The value of water recreation costs in Table 1 were derived in this manner.

**Table 1. United States Annual Pollution Costs, 1970** [1]
(billions of dollars)

|  | Pollutant |  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|---|
|  | Air |  |  |  |  |  | Water |  |
|  | $SO_x$ |  | Particulates |  | $O_x$ |  |  |  |
| Effect | Best | Range | Best | Range | Best | Range | Best | Range |
| Aesthetics and Soiling [2] | 2.9 | 1.7–4.1 | 2.2 | 1.7–4.1 |  |  | 0.1 | 0.03–0.18 |
| Human Health | 1.9 | 0.7–3.1 | 2.7 | 0.9–4.5 |  |  | 0.7 | 0.3 –1.0 |
| Materials | 0.6 | 0.4–0.8 | 0.2 | 0.1–0.3 | 0.9 | 0.5–1.3 | 2.6 | 2.5 –2.7 |
| Vegetation |  |  |  |  | 0.2 | 0.1–0.3 | 0.8 | 0.6 –1.0 |
| Animals |  |  |  |  |  |  |  |  |
| Recreation |  |  |  |  |  |  | 1.7 | 1.2 –2.2 |
| Water Supply |  |  |  |  |  |  |  |  |

[1] Blank cells indicate that values have not been estimated.

[2] Based on property value changes.

Source: Waddell, Thomas E., "The Economic Damages of Air Pollution" unpublished report, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, N.C., February 1974. Tihansky, Dennis P., and Richard Walsh, "A Status Report on the National Benefits of Water Pollution Control" Washington Environmental Research Center, Environmental Protection Agency, Washington, D.C., November 1973.

People's fear of environmental catastrophies or their dissatisfaction with the knowledge that the environment is not being preserved in the best possible quality is determined by a survey which would ask them to make tradeoffs among many different things they want. For example, would they rather have cheaper electricity or more miles of pollution-free water? In practice, these survey techniques have only recently been tried and are only of limited success.

Even though estimating the psychic costs associated with pollution may be very difficult, this does not mean that it is unimportant. In fact, many of the environmental goals and many of the objectives of a comprehensive environmental management plan are derived from the fact that people desire these goals that affect their state of mind and their ability to enjoy the world.

### Avoidance Costs

Avoidance costs, the third kind of costs resulting from pollution, are out-of-pocket expenditures incurred to avoid damage and/or reduce psychic costs.

Examples of this are people incurring extra expense to paint metals so that they will not deteriorate, people moving from a polluted location to an unpolluted one to avoid getting sick, people moving from a polluted environment to a non-polluted one to enjoy the increased aesthetics. It includes the decisions to stop growing crops that are affected by pollution and growing otherwise less profitable crops that are resistant to the pollution. It includes people incurring costs by pretreating water or air so that the pollutants are removed before the medium is consumed. Many of these costs can be estimated if the share of the total expense that is incurred because of pollution can be determined. Only that share should be counted as pollution avoidance costs.

## IV. POLLUTION HEALTH COSTS

Most societies are very concerned with the effects of pollution on the health of its citizens. This effect deserves and has received a good deal of attention. There are five ways that health effects impact upon the citizen, his family and community.

1. Increased doctor, medicine and hospital costs
2. Value of loss from premature deaths
3. Value of time lost from work or other activity
4. Reduced productivity of work if performed
5. Value of suffering from sub-clinical effects

Some or all of the above costs would apply to any given health effect. Some of these components of damages can be estimated in economic terms, although it is difficult. The biggest problem with estimating the value of medical costs, lost activity days and reduced productivity is determining the share of the total observed effect that is attributable to the pollutant when other factors also affect these costs.

The problem of placing an economic value on premature death has often been described as putting a value on life, when it is the length of life that is affected. Determining a value for reduced length of life is very difficult, particularly if age, occupation, health status, and family status obtain different values for individuals. There is no way to determine the correct or true value of an additional day of life, and none will be proposed here. However, it is clear that the value cannot be infinity, which is the value people assign if they insist that a life should be saved (actually increase length of life) regardless of costs. It is also clear that the value is not zero, which would imply that nothing would be done to extend length of life, even if it were very cheap or free. Given the importance societies attach to premature deaths and the difficulty of placing a value on them, it is often necessary to present the number of days lost because of premature deaths to the decision maker and let him make the tradeoff.

Determining the value of suffering is also difficult. In principle, people are willing to put up with some suffering rather than give up goods and services they enjoy. At some point, however, they would be willing to give up some goods and services in return for lower levels of suffering. The value of goods and services people would give up (their willingness to pay) for a given reduction in pollution exposure to reduce suffering would be the correct value to use. In the absence of good estimates of these values, the total number of hours of suffering as a function of pollution exposure should be presented as the damage function.

Because of the difficulty of putting economic values on premature death and suffering, the health damages from pollution can be presented as three damage functions: one showing the number of days lost because of premature deaths, one showing hours of suffering, and the third showing the economic value of all other health impacts. One health damage function that has been estimated is shown in Figure 2. The damage function shows the percentage of a normal, healthy population between ages 10 and 65 experiencing some ill effects as oxidant levels increase. The upper and lower curves are the 95% confidence bounds.

Also, functions of mortality rate versus particulate, sulfate, and socioeconomic variables were estimated with regression analysis, and a significant, positive linear function for each pollutant was obtained [Lave, L. B. and E. P. Seskin. Air Pollution and Human Health. Science *169*(3947):723–733, August 21, 1970]. The number of persons to die per 100,000 population increased 0.041 per microgram

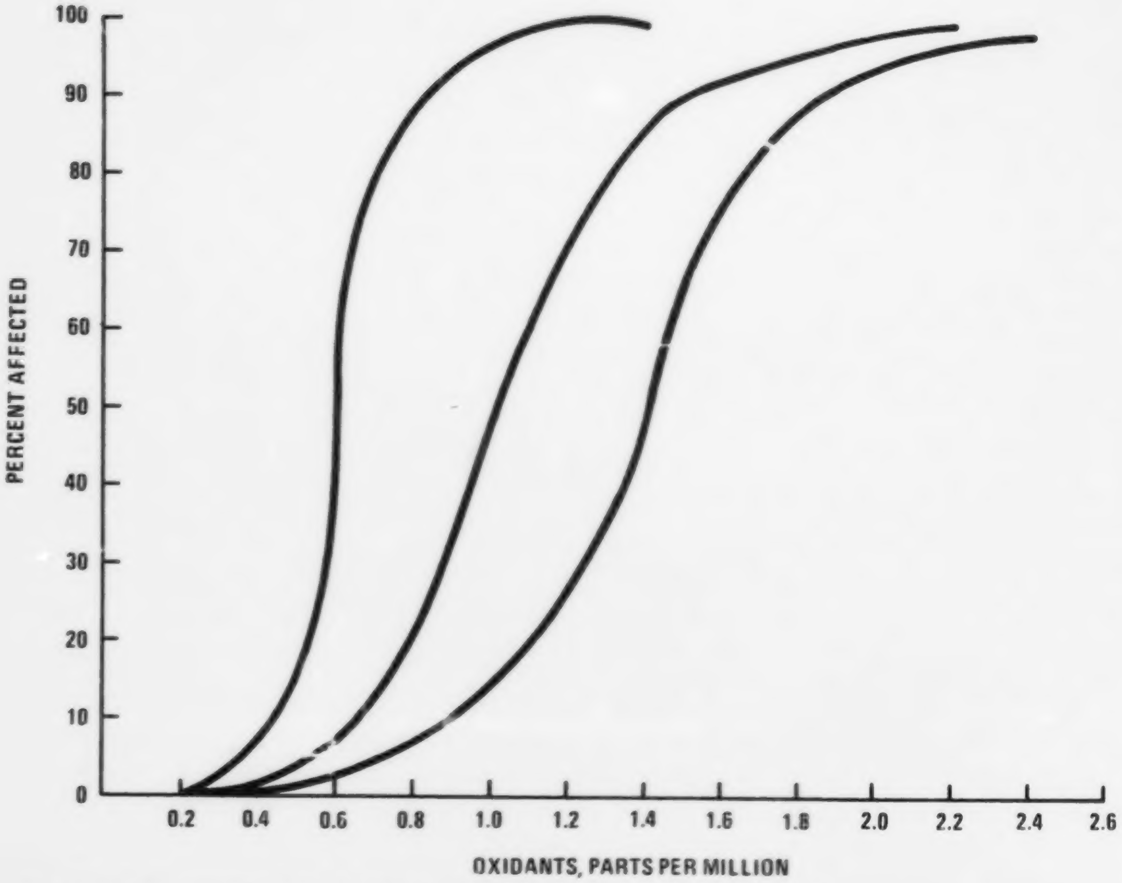


FIGURE 2. PERCENTAGE OF NORMAL POPULATION AGE 10-56 AFFECTED BY OXIDANT EXPOSURE

per cubic meter increase of average annual particulate reading, and 0.071 each microgram per cubic meter increase in sulfate level. A 95% confidence bound for the particulate coefficient is 0.009 to 0.073, and for sulfate coefficient is 0.027 to 0.115. These rates are reliable only for pollution levels near the mean of 4.7 micrograms sulfates and 120 micrograms particulates per cubic meter and with the level of other variables held constant at their means. Including both pollutants in the analysis is superior to estimating the effects of each and ignoring the level of the other. If synergism exists, it can be incorporated by the inclusion of interaction terms. The results would then indicate that the effect of an increase in one pollutant depends upon the level of the other. Thus, the decision about the acceptable level for both pollutants would have to be made at the same time.

Information presented to the decision maker when synergism is present must be different than that discussed above. For example, instead of a single damage function for a pollutant, a set of damage functions must be presented. Each function in the set would assume a different level for the synergistic pollutant.

Health effects damages of water pollution have also been determined. This was done by determining the number of illness outbreaks in the U.S. and determining the share of those attributable to water pollution. A value per illness was then used to obtain the estimate of $0.7 billion per year as shown in Table 1. These estimates are crude and dated. It is expected that more accurate and comprehensive estimates would be several times larger.

## V. COST FUNCTIONS

When a decision is made to control pollution by reducing the level of current emissions, by preventing the increase of emissions or by relocating the source of the emissions or receptors, it imposes costs on society. There are costs of administering and enforcing the decision, costs to plants for add-on equipment, process changes, or changed input mix, as well as costs to society of poor quality products, increased unemployment and reduced economic growth.

These costs can be quantified in economic terms, except for the costs (burdens) to society from decreased employment and reduced economic growth. Thus, it is necessary to supply the decision maker with information on three kinds of adverse effects: employment, economic growth and economic costs.

The costs of obtaining a given level of pollution emission control can vary substantially, depending upon the implementation strategy used. If the time period for obtaining control is lengthened, the costs would be reduced as plants have time to phase in new equipment and because new, more efficient technology may be developed. If control applies to all plants, only large plants, only new plants, etc., costs will vary widely. Thus, cost analysis requires that the implementation strategy be pre-determined, several alternatives be studied or the least-cost (economically most efficient) method be determined. In the least-cost case, the required level of control is based on the cost of control. The optimum solution has the cost of an additional unit of control equal for all plants. In all cases, it will be necessary to determine the number, distribution and kinds of sources, the control options and their associated costs available to each source, and the probability that the control costs would be so high as to force the closing of the plant.

Studies have shown that the least-cost strategy has significantly lower costs than strategies designed to impact each source uniformly. A least-cost solution is obtained by linear programming, where each kind of source is considered, available control options are included as activities and total emissions are the constraints. The solution specifies the control option and emission level for each kind of source such that total costs are a minimum. Alternative solutions are obtained as the constraints (total emissions) are varied and a least-cost cost function is derived.

Cost functions have been estimated for selected regions. The curves in Figure 3 show the difference between the least-cost and a uniform plant strategy that was being implemented. Even though the least-cost strategy is economically most efficient, there may be political or institutional constraints that prevent it from being used. Nevertheless, least-cost functions should be derived in all studies so the decision maker realizes what is given up to accommodate these other factors.

## VI. COMPARABILITY OF COSTS AND BENEFITS

As stated earlier, when a decision is made, the advantages and disadvantages of that decision and alternatives to that decision have been compared, either explicitly or implicitly. In this process of comparing advantages (benefits) and disadvantages (costs), it is useful if they are measured in the same units. Although many different units can be used, monetary units have many advantages. One advantage is that many of the effects are measured in monetary units such as the costs of add-on control equipment and medical expenses from pollution damage. Another is that the monetary units are easily understood. The difficulty of measuring all the adverse and beneficial effects in economic terms has been discussed. In practice, it may be necessary and desirable to present the decision maker with several functions, measured in different units, i.e., number of premature deaths, hours of suffering, reduced employment, etc. Here it is assumed that they can all be measured in economic (or other unit) terms so that determining the socially desirable level of pollution can be explained. Later, this assumption will be relaxed and the process of making trade-off decisions when effects are measured in different units will be explored.

In making decisions about the acceptable level of control, the change in damages must be compared to the changes in costs for a proposed change in ambient quality. These changes are called marginal damages and marginal costs. Marginal damage functions and the negative of the marginal cost functions with their associated variance are presented in Figure 4. The negative of the cost function is used to get both marginal functions in a single quadrant. Both sets of data, the costs and the benefits, have been annualized so the figure depicts costs and damages per year. Point $P_1$ in Figure 5 represents the socially optimal level of pollution control. Except for very unusual situations, that level of control will not be zero emissions, zero effects, or uncontrolled emissions, zero effects, or uncontrolled emissions. At Point $P_1$ the marginal costs are just equal to the marginal damages. It can be demonstrated that the point where the marginals are equal is the optimum level. The procedure for determining the optimum point is not different for a damage function with a threshold like function B in Figure 1.

The A', A'' and C', C'' functions reflect the fact that neither the cost function nor the damage function is known with certainty. The shaded area in Figure 4 shows the area within the confidence bounds of both functions. Because of uncertainty, any pollution level corresponding to a point in the shaded area could be the optimum level. Thus, the optimum level could, with the confidence implied by the confidence bounds, be somewhere between $P_2$ and $P_3$. The more uncertainty we have about the damage and cost functions, the larger will be the range that contains the optimum level and the more

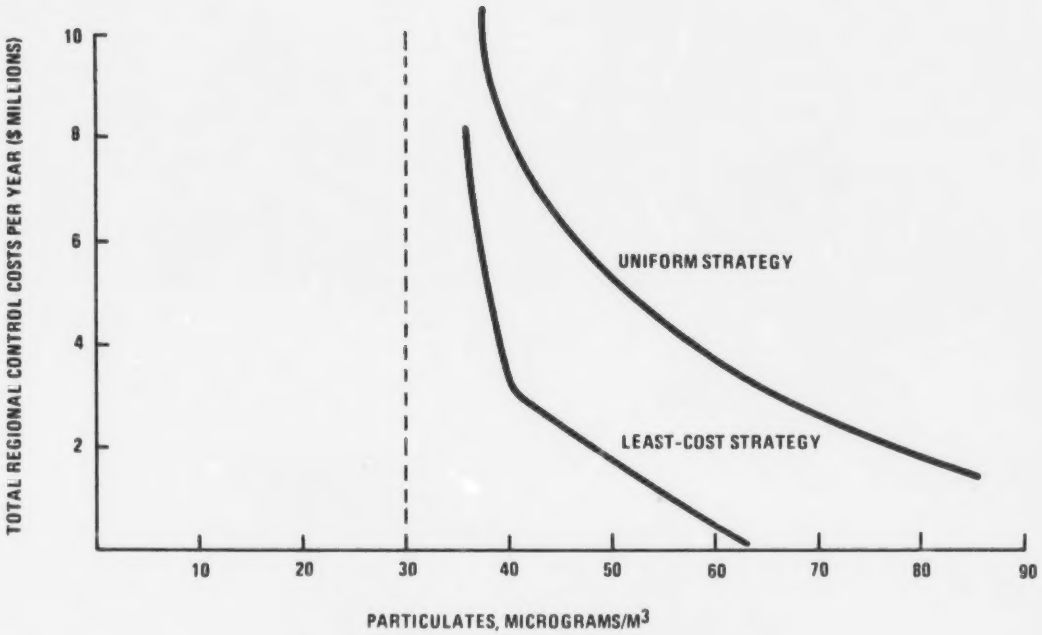


FIGURE 3. COST FUNCTIONS

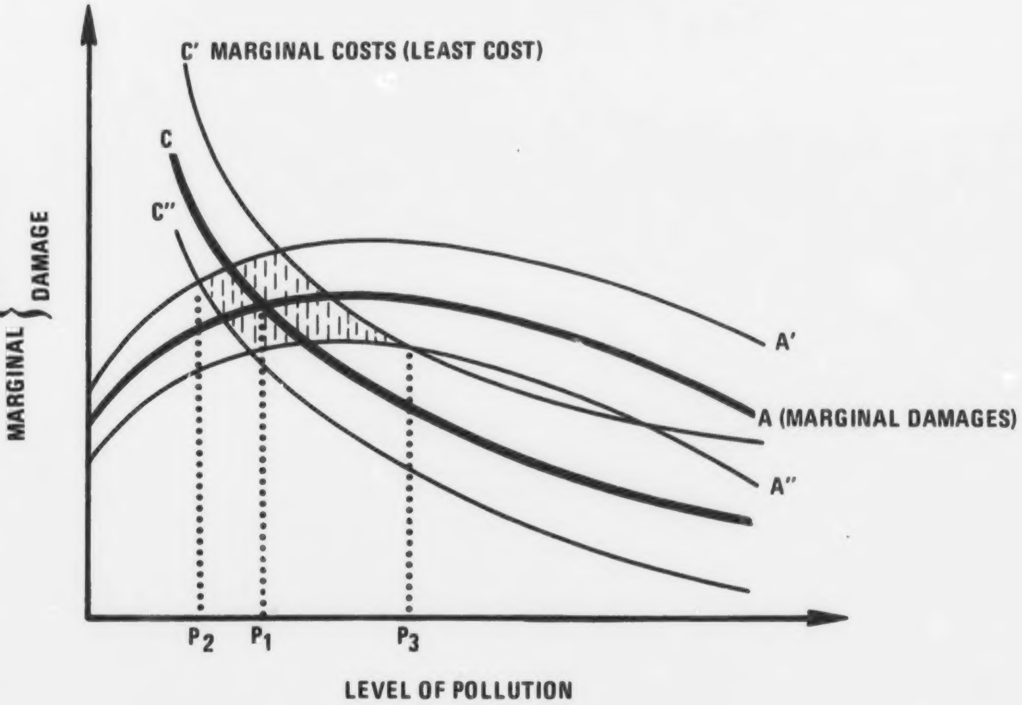


FIGURE 4. MARGINAL DAMAGE AND MARGINAL COST FUNCTIONS OF CHANGING THE LEVEL OF POLLUTION

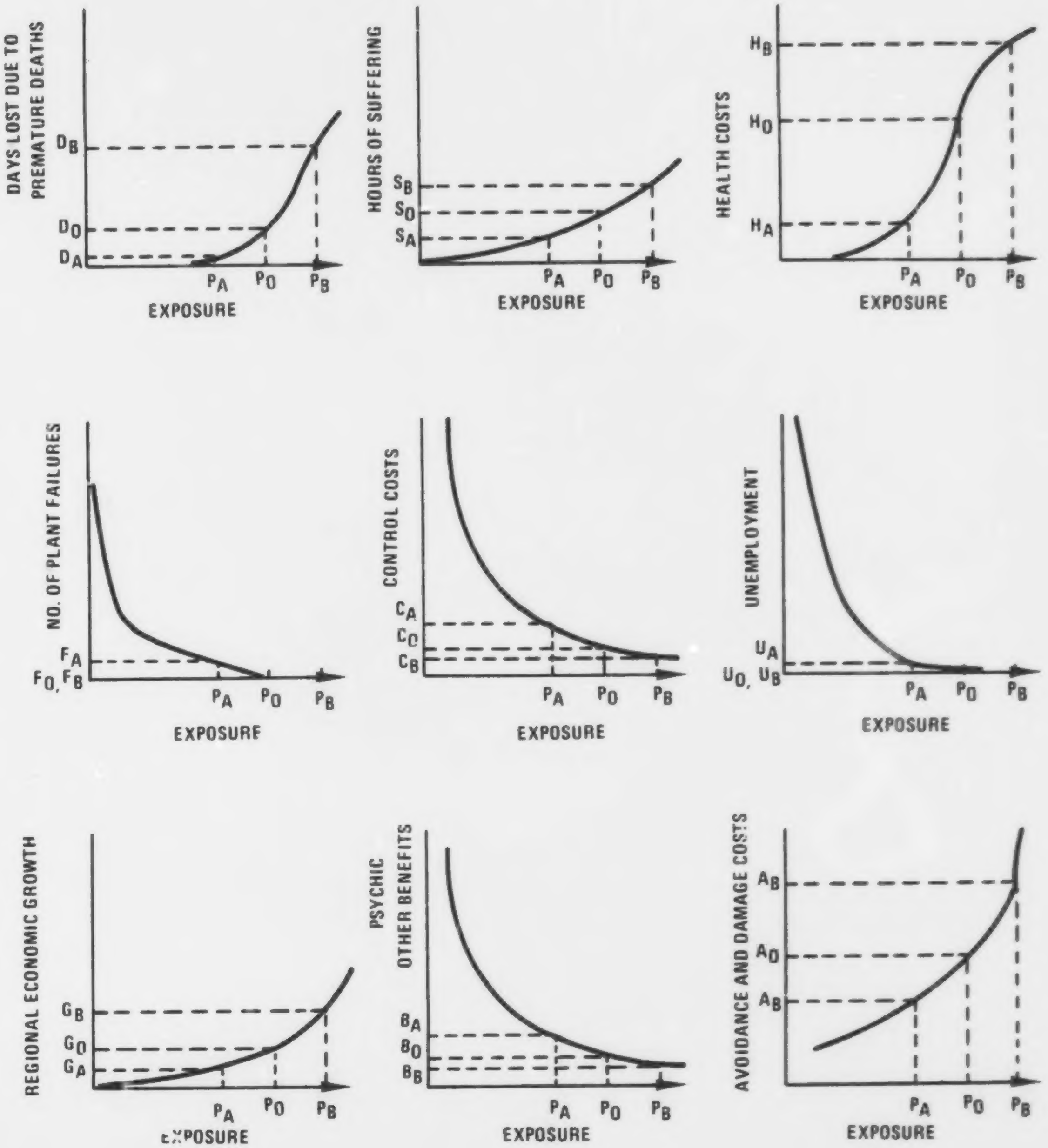


FIGURE 5. DAMAGE AND COST FUNCTIONS

difficult for decision makers to choose an acceptable level of pollution with confidence that it is near the correct level.

## VII. DECISION MAKING WITH MULTIPLE BENEFIT AND COST FUNCTIONS

The current state-of-the-art in estimating both costs and benefits does not allow the ideal of getting all the effects collapsed into a single functional form so that the optimum decision point can be easily determined. In fact, many of the effects simply cannot be measured at all. As stated earlier, many of the costs, particularly when emission control is obtained by relocation or by changing the process or the product quality, also have not been measured. So, the decision maker has to settle for estimates of some key effects that are not in comparable units.

For health damages there are three key relationships: number of years/days lost because of premature deaths, the total health costs including doctor and hospital costs and lost days of work, and the hours of suffering from sub-clinical effects, each as a function of air pollution levels. Four key variables on costs are the total costs of add-on equipment with the associated operating and maintenance costs, the number of plants forced out of production, the decreased employment in the region, and the foregone benefits of economic growth. The two key variables for environmental pollution costs are psychic costs and economic value of avoidance and damage costs. Hypothetical functions for these nine functions are shown in Figure 5.

The level of exposure $P_o$ results from the existing level of pollution. This becomes the base-line case for the analysis. Decision makers could use a set of functions as depicted in Figure 5 to make their own trade-off decisions. However, there are many tradeoffs involved, and it may be easier if selected results were displayed in a table. Results for several alternative control levels can be presented in such a table. Examples are shown in Table 2. Alternatives A and B are shown. The values in the cells are the changes from the base case.

Alternative A considers a reduction in exposure levels. Thus, in comparing it with the base case, the tradeoffs are fewer premature deaths, (Do-Da), fewer hours of suffering, (So-Sa), lower health costs, (Ho-Ha), increased psychic benefits (Bo-Ba) and reduced avoidance and damage costs (Ao-Aa); traded off for higher control costs (Co-Ca), reduced regional growth (Go-Ga), increased plant failures (Fo-Fa), and increased unemployment (Uo-Ua). A decision maker would have to weigh these different effects to determine overall costs and benefits of Alternative A.

Alternative B presents similar tradeoffs from allowing pollution levels to be worse. Many alternatives could be presented to the decision maker in this format. Also, the number of effects can be increased to include important subsets. If it were important to the decision maker to know how premature deaths were distributed between young and old, separate functions would be generated and separate rows presented in the table.

The procedure described above demonstrates how information for cost-benefits (trade-off) analysis can be presented to decision makers when there is a problem of weighing the different effects. It assumes that functions are available for these effects. In fact, very few functions have been estimated.

## VIII. STATUS OF COST-BENEFIT ANALYSIS IN THE UNITED STATES

Cost-benefit analysis has been used for many years as a decision tool for water resource projects. Much of the current state-of-the-art of cost-benefit analysis for environmental objectives has come from this experience. However, environmental cost-benefit analysis requires estimates of many benefits not important to water resource decisions.

Cost-benefit analysis is required in many of the laws governing the establishment and enforcement of acceptable levels of control. However, requirements for its use vary considerably.

In the laws governing the primary ambient level of air pollutants, the level is to be determined by threshold health effects, and costs and benefits are not to be considered. As it becomes clear that thresholds do not exist or are very low, consideration is being given to shifting to a cost-benefit criterion for determining levels.

Special laws were passed requiring emissions from mobile sources to be rolled back 90%, with limited

**Table 2. Selected Results From Damage and Cost Functions**

| Effects | Base | *Alternative Control Targets* | |
|---|---|---|---|
| | | A | B |
| Number of Premature Deaths | $D_O$ | $D_O-D_A$ | $D_O-D_B$ |
| Hours of Suffering | $S_O$ | $S_O-S_A$ | $S_O-S_B$ |
| Health Costs ($) | $H_O$ | $H_O-H_A$ | $H_O-H_B$ |
| Number of Plant Failures | $F_O$ | $F_O-F_A$ | $F_O-F_B$ |
| Control Costs ($) | $C_O$ | $C_O-C_A$ | $C_O-C_B$ |
| Unemployment | $U_O$ | $U_O-U_A$ | $U_O-U_B$ |
| Regional Growth | $G_O$ | $G_O-G_A$ | $G_O-G_B$ |
| Other Benefits | $B_O$ | $B_O-B_A$ | $B_O-B_B$ |

consideration of the costs or benefits in setting this level and date of attainment. Similarly, very stringent transportation controls are being relaxed and postponed, in cases where it has become apparent that the social costs would be very high.

The laws considering water quality vary in their requirements for cost-benefit analysis. All plans by states to implement water pollution programs are required to be evaluated in a cost-benefit analysis and others require such analyses to establish effluent discharge limits. If individual firms believe the costs of complying with the limits exceed the benefits, they can request a cost-benefit analysis; and if costs exceed benefits, they do not have to comply at that limit.

Limits on toxic substance emissions to the environment are to be established on the basis of cost-benefit analysis. Determination of whether a pesticide can be used has to be made after consideration of the costs and benefits.

Although requirements for cost-benefit analysis vary among pollutants and change over time, increasing consideration is being given to using it more in making decisions. The need for improved methods and improved data, particularly dose-response relationships, is increasing rapidly.

# Methods for Abatement of Water Pollution

Robert B. Schaffer *

## I. INTRODUCTION

The comprehensive environmental analysis of a defined area requires the assessment of all sources of pollution, their interrelationships and ultimate control possibilities. This paper has two objectives:

1. to provide the participants of this symposium necessary information on water pollution sources and associated control strategies; and
2. to look at some of the more conceptual strategies that may be termed Operational Control.

The former information includes the description of levels of control and their cost for some representative industrial categories and the municipal or domestic waste source. The diffuse or non-point sources cannot be as precisely scoped. The information, therefore, will only provide data for decision-making and not make the decisions. The strategies employed in Operational Control must integrate the technologies described and incorporate them in the decision-making process through the development of basin or region plans. Specific data on the relative costs and socio-economic benefits have not been included, but they must be weighed with all other aspects of the overall environmental plan. Many social and legal/regulatory factors must also be considered, and some of the obvious will be briefly discussed.

## II. POINT SOURCES

Point sources of pollution from municipalities, industry, commerce, and agriculture are the most controllable and best lend themselves to the improvement of water quality in a cost effective manner. Technology is available to achieve the reduction of most pollutants, to the limit of detectability, if cost is no object. However, this option is not generally available, and other alternatives must be considered.

The single variable that provides the most potential for achieving the greatest removal from point sources is time. A long-range strategy must be developed for the region to achieve maximum control at the lowest cost. The objective is to do the best job you can economically, not to get by as inexpensively as possible, which is the tendency when opposing interests attack a common problem.

The optimum point source control strategy then should have as its goal zero discharge of pollutants, defined as the removal of undesirable constituents of wastewater to levels of detectability. Zero discharge does not mean the removal of all constituents nor the elimination of the liquid discharge itself. In these instances primarily, but in all others as well, the pollutants so removed from discharges to waterways become materials of concern in other environmental areas. They may become air pollution problems through incineration, solid waste problems through conversion and disposal to the land, or a problem due to their presence in products produced for consumption and use. A water-borne discharge may turn out to be the best ultimate disposal option for the overall environment.

The most complex and difficult problems in development of an effective environmental assessment of an area are associated with an industrialized society. Treatment or abatement technologies for the varied and necessarily complex manufacturing processes stretch the limits of existing technology when viewed in the context of zero pollution. Whether municipal or industrial, a certain amount of that which results from man's industrial society must be tolerated.

One basic, but often forgotten, law in relation to pollution control is the conservation of matter. Many would have those materials that present environmental problems just disappear. Unfortunately, they will not, no matter how hard we try. There is no such thing as infinite recycle or reuse. Therefore, difficult technical, social and economic decisions must be made with regard to each specific problem.

The goal of zero discharge must be approached, therefore, in a step-wise fashion; Public Law 92-500 outlines the pragmatic, technology-based approach now underway in the United States.

1. The application of Best Practicable Control Technology Currently Available to all existing point sources other than publicly owned

---

*Permit Assistance and Evaluation Division, U.S. Environmental Protection Agency treatment works, i.e., the average of the best practiced technologies to be in operation by July 1977;

2. The application of the Best Available Technology Economically Achievable, i.e., the best that could be built using possible scale-up of pilot technology; to be in operation by 1983;
3. The application of the Best Demonstrated Technology, including use of different manufacturing processes, to all new sources;
4. The application of pretreatment requirements, which must control incompatible pollutants, on discharges to municipal treatment systems;
5. The application of secondary treatment to all existing municipal systems, this step on the same time frame as Step 1 for other types of discharges;
6. The application of Best Available Technology to municipal treatment works, this step to coincide with Steps 2 and 3.

The application of this strategy is complicated by the diversity of industrial categories and the ability of technology and economics to cope with these problems. In some categories it will be impossible to achieve complete recycle and may well be undesirable. This paper examines, for purposes of illustration, the application and costs of various levels of pollution abatement in different industrial and municipal situations. Both major industrial segments shown, Iron and Steel and Petroleum Refining, do not provide for the goal of zero discharge of pollutants. Petroleum refining effluents can be controlled more easily, primarily due to the technical sophistication associated with the processing of the raw materials into the various products. The Iron and Steel Industry requires high volumes of water for cooling and quenching and presents a less controllable option, from a water recycle standpoint. The Metal Finishing category presents the "no discharge" option, primarily because of its size and product recovery economics. In this instance, Step 2 (Best Available Technology) will be equal to Step 3 (New Source Requirements) which is zero discharge.

The treatment of municipal wastes using existing technologies takes us up to the renovation of water for beneficial reuse. Effluent quality may be tailored to the use required, from water quality maintenance through various stages of water reuse.

## Petroleum Refining

A petroleum refinery is a complex combination of interdependent operations, and the characteristics of the wastewater differ considerably for different processes. Each process is itself a series of unit operations which causes chemical and/or physical changes in the feedstock or products, and each operation may have drastically different water usages associated with it. Table 1 shows the discharge concentration averages for both Best Practicable and Best Available levels of treatment. The difference in the amount of pollution discharged in this case is related to a significant reduction in the volume of water used per unit of production for Best Available Treatment.

## Iron and Steel Industry

The industry is comprised of a complex of manufacturing processes and wastewater treatment operations, both biological and physical-chemical, which may be used. The proposed Best Available levels for the Iron and Steel industry indicate that it requires and will continue to require large volumes of water for the manufacturing process. Therefore, in long-range planning, this consideration must be integrated into the ultimate location of the manufacturing facilities. Tables 2 and 3 show the discharge concentration averages for the industry for Best Practicable and Best Available levels.

**Table 1. Petroleum Refining Industry**

Attainable Concentrations from the Application of
Best Practicable Control Technology Currently Available
and
Best Available Treatment

| Parameter | Annual Average (mg/l) |
|---|---|
| Biochemical Oxygen Demand | 15. |
| Chemical Oxygen Demand | 80. |
| Suspended Solids | 10. |
| Oil and Grease | 5. |
| Phenol | 0.1 |
| Ammonia Nitrogen | 80% Removal |
| Chromium, Total | 0.25 |
| Hexavalent Chromium | 0.005 |
| Zinc | 0.5 |

**Table 2. Iron and Steel Industry**

Attainable Concentrations from Application of
Best Practicable Control Technology Currently Available

| Parameter | Concentration (mg/l)* |
|---|---|
| Biochemical Oxygen Demand ($BOD_5$) | 150. |
| Suspended Solids | 50. |
| Oil and Grease | 10. |
| Phenol | 3. |
| Ammonia (as Nitrogen) | 125. |
| Cyanide, Amenable to Chlorination | 7. |
| Chromium, Total Dissolved | 0.05 |
| Iron, Total Dissolved | 1.2 |
| Zinc, Total Dissolved | 1.2 |

*With Recycle

## Metal Finishing

The basic treatment model is that of unit process stream precipitation. For essentially all of the parameters, Best Practicable involves precipitation, which includes coagulation, sedimentation, flotation, and

**Table 3. Iron and Steel Industry**

Attainable Concentrations for Application of
Best Available Control Technology Economically Available

| Parameter | Concentration (mg/l)* |
|---|---|
| Biochemical Oxygen Demand | 20. |
| Suspended Solids | 10. |
| Oil and Grease | 10. |
| Phenol | 0.5 |
| Ammonia (as Nitrogen) | 10. |
| Cyanide, Amenable to Chlorination | 0.25 |
| Chromium, Total Dissolved | 0.05 |
| Iron, Total Dissolved | 1.2 |
| Zinc, Total Dissolved | 1.2 |

*Treatment of Recycle Blowdown

finally, filtration. Evaporation recovery and ion exchange offer other effective approaches, and in some cases, chemical oxidation (for cyanides) and chemical reduction (for chromium) are needed as part of the treatment process.

The costs of achieving effluent reduction to "no discharge" have been examined for the industry. The cost may be about the same as for conventional chemical treatment because it may be possible to recover and reuse sufficient chemicals and water to offset the costs of treatment.

## Municipal Wastewater Treatment

In making determinations with regard to handling of domestic wastewaters in the context of area planning, certain considerations become more important than others First, the combination of compatible industrial wastes may present economical solutions and must be considered. Secondly, the existence of defined areas of development may limit the choice available, usually on social and economic grounds. Thirdly, where new systems may be designed, there are fewer constraints on renovating water for reuse.

Of immediate interest in the handling of wastewater management problems in areas already under development is the upgrading of present treatment plants. Technology in the past decade has provided many innovative upgrading procedures to meet deficiencies in existing processes. Performance and cost data on the upgrading of facilities have been developed.

The consideration of advanced waste treatment processes for water renovation and potential reuse can result in a viable, economical alternative. Suhr has described cost and performance data for the South Lake Tahoe Public Utility District. The results of treatment meet or exceed the stringent requirements imposed. These results are summarized in Table 4.

All areas do not require the same degree of waste treatment, and all of the tertiary processes utilized at Lake Tahoe are not required at each proposed tertiary treatment plant. Because they are unit chemical processes, the Tahoe processes can be adapted to provide treatment to meet almost any practicable pollution abatement criteria.

**Table 4. Reclaimed Water Quality**

First Year of Plant Operation
South Tahoe Public Utility District

| *Quality Parameter* | *Median* | *Maximum* |
|---|---|---|
| BOD mg/l | 1.0 | 7.8 |
| COD mg/l | 10.8 | 24.7 |
| Phosphate (as Phosphorus) mg/l | 0.42 | 4.8 |
| Jackson Turbidity Units | 0.6 | 6.0 |
| Suspended Solids mg/l | 0.5 | 7.8 |
| Coliform MPN/100 ml | 2.2 | No two consecutive samples greater than 16. |

The overall consideration of point source control is summarized in Figure 1. Considerations for the implementation of pollution abatement beyond Best Practicable, which is assumed to be the base acceptable level, must consist of an area's water use requirements, volume available, water quality, and the assimilative capacity of these waters to maintain desired quality. In general, the costs for achieving the next level, that is, Best Available Treatment, will be roughly three times the cost of Best Practicable.

To go beyond, to approach zero discharge of pollutants, will increase the cost significantly. The desirability of this option, where technically feasible, will depend directly on the value of the water in that specific location.

## III. NON-POINT SOURCES

### Characteristics

The fact that nearly all non-point sources are rainfall/runoff dependent and are associated with highly variable natural phenomena has profound implications for their control. It is extremely difficult to characterize a non-point source or evaluate control performance dependent on such variables as rainfall and growing season which cannot be predetermined, controlled or accelerated. Accurate assessment of the problem may require a significant period of time (one to three years). Because there is no typical year or season, a multiyear study may be required.

Non-point sources have enormous diversity. Discharges are noncontinuous and highly dependent on location and local conditions. No single set of conditions would be representative, and an extrapolation of data from specific studies is risky. These sources have no defined discharge point and cannot be conveniently or economically measured. The discharge or load must be related to some measurable source feature or activity. Because control by effluent treatment is impractical in most instances, we must control the activity by causing a change in the pollutant's manufacture or use. Source control problems involve significant institutional, legal, and socio-economic interfaces.

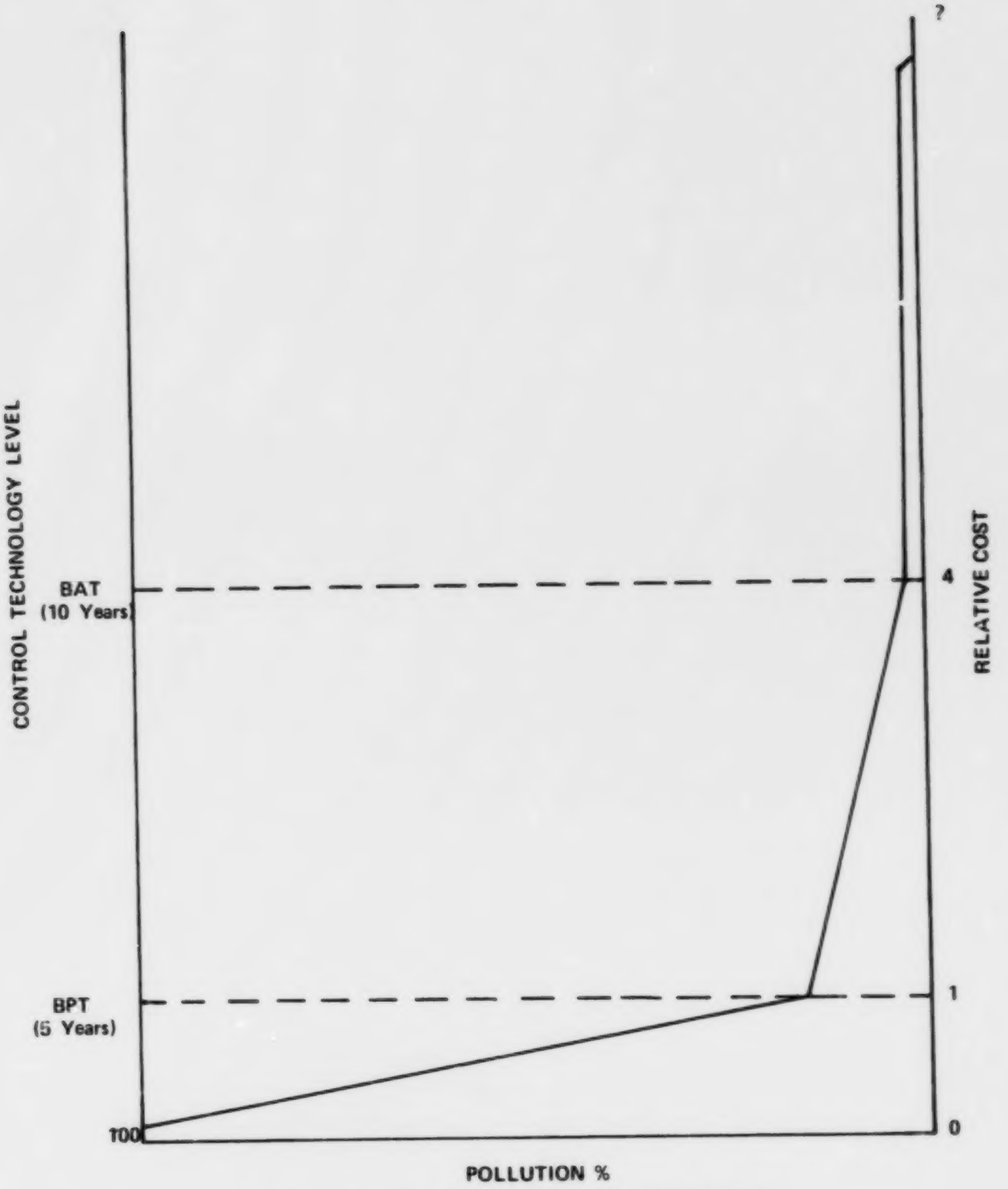


FIGURE 1. RELATIVE COST OF IMPROVING WATER POLLUTION CONTROL

## Proposed Control Strategies

Relief from non-point source pollution requires, as a first step, the assessment of the problem in order to set proper program priorities. The leading models for non-point sources and representative land use patterns should be developed. The probably effectiveness of available control options must be assessed. Basin load allocation models must be developed to allow trade-offs in individual situations between the control of point sources and non-point sources.

The following is a brief discussion of six specific non-point source areas.

## Irrigation Return Flows

This water quality problem involves all irrigated agriculture, including water and crop management procedures and accompanying use of chemicals. Major pollutants from return flows include salts, sediments, nutrients, toxic and other bioactive materials. The magnitude of the problem is indicated by the fact that approximately 10 percent of the cultivated acreage in the United States is under irrigation, producing 25 percent of the national food and fiber. Doubling of irrigated acreage by the year 2000 with a 10–20 percent increase in total water use is projected. With existing practices, total salt discharged could double. Salt has extensive effects on agriculture to the extreme of rendering farm land no longer productive. Fifty to seventy percent of man-induced salt sources in receiving waters can come from irrigation return flows. This pollutant will present a major interbasin problem and can affect international uses of water. Control strategies must consider soil and water management practices in each climatic/geographic location and the strategy, therefore, will vary with location. Source management is the most cost-effective general control option now available, entailing water management as well as the management of the application of various chemicals used in irrigated agriculture.

### Animal Feeding

All animal raising operations produce pollution from the oxygen-demanding materials, nutrients, salts, and pathogenic organisms present in waste waters from the operations, with significant impact in any basin where such activity is a major factor. Present trends relate to the economics of larger operations. In the United States sixty percent of the marketed beef cattle are produced by fewer than two percent of the operations. Control strategies for many operations include land disposal of residues produced, and containment and spray irrigation with water resulting from such operations. The economics of hauling wastes, and climate are problems here. An ultimate solution is likely to be by-product recovery and land disposal of effluents.

### Agriculture Runoff

This segment includes crop raising, animal pasturing, forestry and other rural land uses from all nonirrigated agricultural operations. Major pollutants include sediments, nutrients, oxygen-demanding materials, organic and inorganic residues from agricultural chemicals, toxic and other bioactive materials, heavy metals, and pathogens. This activity in any given basin may cover at least fifty percent of the land area, resulting in significant amounts of pollution due to runoff from precipitation. The trend towards sharply increased agricultural production both in acreage and in productivity is likely to accelerate pollution by use of marginal lands and even greater amounts of chemicals. Control strategies focus primarily on source control management, but highly variable and complex conditions make control of this diverse situation very difficult.

### Hazardous Material Spills

Hazardous materials production, storage, transportation, or utilization in close proximity to a water source provide a potential hazard to water quality. Major pollutants from spills are toxic and other bioactive chemicals, acids, alkalies, corrosives, oxygen-demanding materials, and petroleum oils and related products. The significance of the problem is manifest, but its magnitude is difficult to assess. Control requires detailed, comprehensive planning and action response capabilities to minimize harmful effects.

### Mining

All surface and underground mining activities, including milling, benefication, and other site operations, are sources of pollution. Major pollutants include sediments, acids, toxic metals, and other salts. Increased demands for energy and all types of minerals will result in increased impact of active mining operations. Long-term problems will continue as our requirements continue and increase. Control strategies, depending on the source, include surface reclamation and specific collection and treatment of discharges. Coal mine drainage treatment, for instance, is costly but not unreasonably expensive. An associated problem is pollution from abandoned mines. Nonoperating surface and underground mines, including refuse and tailing piles, provide significant pollutant amounts. Control is often difficult because specific sources cannot be identified.

### Land Modification

All of man's activities which alter land surface and thereby facilitate runoff (construction, landfilling, etc.) produce significant amounts of pollutants, primarily sediment carried by surface runoff. The impact of these activities is a universal problem related to overall area economic development. Sediment is both a physical pollutant and a carrier for other pollutants. Source controls are the only apparent promising operations for the control of runoff from construction. Conditions vary widely, but they must be seriously considered in the development of any type of regional plan that expects to significantly modify existing land use patterns.

## IV. OPERATIONAL CONTROL

Wastewater treatment technology describes alternatives available to achieve a desired result through its application to point sources. Non-point sources may limit the potential for achieving all desired water quality goals through point source control alone. Other avenues for the achievement of the optimal water quality must be assessed, and these fall into the area of operational control. Some of these alternates and how they may be incorporated into the overall plan are examined here.

## Non-Point Source Alternates

Since there are no easily or rapidly achievable means of overcoming non-point source pollution, it seems most important to incorporate some basic principles into the long-range objectives. Some of these basic principles or concepts are:

1. the establishment of land use requirements to minimize the concentration of industry and other associated development;
2. modification of agricultural practices to minimize runoff and the associated pollutional problems of silt, fertilizers, pesticides, herbicides, and organics; and
3. the control of area runoff (stormwater) which can be as much of a pollutional load from urbanized areas as the wastewater produced.

## Land Use

The establishment of land use requirements may well be the most difficult and certainly most complex problem facing an areawide planner. An optimum system for one purpose, water pollution control, is not adequate. Other environmental concerns, such as air pollution control and solid waste disposal, are important as well as considerations of traffic flow, recreational and aesthetic requirements. Other less obvious factors are the location of natural resources, i.e., one must locate a mine where the minerals are, transport the product to the market, supply resources, and provide support functions.

The benefits of establishing a good land use plan from a water quality standpoint are tremendous. Minimizing the stress of pollutants on water resources through selective locations will give this part of the ecosystem a chance to stabilize, that is, reduce the rate of degradation.

## Agripractice

Because the choice of location for agricultural activity is normally limited in any given basin by the availability of suitable lands, the potential for operational controls becomes more important. Agricultural non-point source water pollutants come in large part from natural runoff. Pollution prevention must be achieved through the management of agricultural practices in cropping pasture, forestry and other rural land uses. The basic problem, to manage for proper operational changes to minimize pollution without decreasing productivity, is further complicated by such uncontrollable natural factors as climate, topography, etc. Nevertheless, significant progress can be made. Contour farming methods and the proper application of fertilizers and other chemicals can materially reduce pollutants reaching the water. As in all operational control areas management models can effectively be used for characterization and load reduction predictions. The degree of control attainable will be somewhat less than that for other sources because of the diverse nature of the problem. Even so, the potential water quality gains, especially in areas affected by toxics, eutrophication, and silting, can be well worth the effort.

## Area Runoff

Area runoff, water resulting from rainfall contaminated by contact with man's urbanized and suburbanized activity, results in more diverse pollutants than other non-point sources. Street runoff (containing oil, motor vehicle exhaust chemicals), park and lawn area runoff (fertilizers and toxic organics), and construction site runoff (silt from disturbed areas) will equal or exceed pollution from municipal wastewaters. The volume of water can be extremely high, depending on the location and degree of coverage (e.g., concrete).

Flows of this type collected in storm or combined sewers present a significant point source of pollution that is rarely effectively controlled. The only practiced technology on these large volume point sources is brief settling and skimming. Lower flows or the "first flush," can be routed to municipal or industrial treatment plants.

Manufacturing facility runoff adds to the complexity of the wastes, with material spills, and with runoff and/or leaching of storage piles capable of contributing extremely high pollution loads. Available control methods include management of flow (e.g., dikes), routing of certain types of flows (e.g., roof drains), and covering of refuse or storage piles, etc.

Combined sewer separation, and in developing areas, the construction of separate storm drains and sanitary sewers, is the best solution from a pollution standpoint. The economics of separation in urbanized/industrialized areas may be prohibitive.

Effective operational control of non-point sources will require time and good management. Some of the management techniques will include long-range planning, turning non-point sources into more controllable point sources, and significant predictive and management modeling.

## Water Management

Optimum water management in urban areas requires the serious consideration of water renovation and reuse, particularly in water-short areas. The supply of water is static geographically, unless diverted by man. Will man move to water or develop the techniques to supply water where he lives? Traditionally, man has remained relatively immobile. Pollution patterns in 2070 will probably be similar to today's. The total supply of water is about the same as it was when the earth was very young and it will be the same tomorrow!

The ultimate in recycle or "closing the loop" is recycle of public wastewater directly back to the supply. Psychological factors inhibit this type of reuse despite technical capabilities. Reuse of another city's effluents by a downstream city doesn't seem to bother us as much, though. In other words, the use of another's treated sewage is acceptable, whereas the use of one's own treated sewage is unacceptable. Advanced water purification technology destroys the "logic" of this traditional attitude. A clear, tasteless, odorless water can be produced. This "reuse reluctance" has caused many cities to go to great expense for alternative water sources, such as building large reservoirs and transporting water hundreds of miles.

When faced with a decision on where to go for increased water supply, the existence of these many alternative schemes makes selection of the best one extremely difficult. Many technical, economic, and policy questions must be answered. Actual conditions may cause many of the options to "drop out," making the best alternative obvious by inspection.

## V. NON-TECHNICAL CONSIDERATIONS

A brief mention of some important non-technical aspects of integrating water pollution control into an environment plan is an appropriate way to conclude this portion of the overall assessment. Probably one of the most frustrating problems may be termed legal or jurisdictional. Conflicts will arise, for instance, when one part of an optimum water system must suffer due to the need of a transportation system or vice versa. Bureaucratic battles can be predicted when one local jurisdiction must, for instance, become the repository for the waste materials from another jurisdiction.

Depending on the existing form of government, personal property rights, resources rights, water laws, and industrial interests can serve to delay and change the optimum plan.

Consequent legal challenges are as much of a stumbling block to the orderly progression of a plan as any technical problem.

## VI. CONCLUSIONS

Technology is available to make great strides toward improving water quality in a cost effective manner through point source control. Methods for the control of non-point sources provide less of an opportunity to achieve control in the immediate future and will require continued research and development to develop and perfect control measures. The longer range goal of zero discharge of pollutants will be dependent upon the development of new and more sophisticated pollution abatement techniques.

# Methods for the Control of Air Pollution

K. E. Yeager *

## I. INTRODUCTION

The control of air pollutant impacts is a complex problem because of the variety of sources and source characteristics. Technical and economic factors frequently make it necessary to use different control procedures for different types of sources. Many emission control techniques are still in the developmental stage, and practical near-term control strategies may call for the use of interim methods until these techniques are perfected. In these situations, alternatives, such as the use of design and siting criteria which distribute rather than limit emissions, may be required as a means for achieving air quality goals. Therefore, fundamental to the development of either Air Quality Management or widespread source control strategies for the achievement of air quality goals is an understanding of the performance, cost, applicability and availability of potential control alternatives. This paper will consider these basic factors as they bear on both technical and operational emission control as well as emission dispersion methods.

These considerations will begin with a discussion of the air emission inventories associated with stationary and mobile source categories and their relationship to ambient air quality and proceed to discuss the status of control methods for (1) sulfur oxides, (2) total suspended particulate, (3) fine suspended particulate, (4) nitrogen oxides, (5) hydrocarbons and oxidants, and (6) carbon monoxide.

## II. CONTRIBUTION OF SOURCES TO EMISSIONS

Table 1 summarizes the contribution of source categories contributing at least one percent to the inventory of six major air pollutants in the United States. With the exception of fine suspended particulate, health related ambient air quality standards have been established for each of these pollutants. Fine suspended particulate (less than 3 microns) on the other hand has been shown to be potentially the most offensive particulate size fraction because it is more easily respirable and more readily retained in the lungs as well as having the ability to persist for extended periods in the atmosphere.

### Sulfur Oxides

Over 60 percent of the total sulfur oxides emissions result from the combustion of coal, primarily in large electric generating facilities. About two-thirds of the nationwide emission occurs in urban areas with power plants, industries, businesses and homes all contributing to the burden. Much of the pollution in rural areas is emitted by a relatively small number of large individual sources. For example, individual smelters, generally located away from metropolitan areas, may emit over 100,000 tons per year. In addition, almost half of the Nation's power plants, including seven of the eight largest sulphur oxide emitters, are located in rural areas. Annual emissions from these seven range from 200,000 to 300,000 tons per year.

### Nitrogen Oxide

Motor vehicles and the statutory combustion of fossil fuels account for about 90 percent of the national total. Nitrogen oxides and carbon monoxide, probably more than other pollutants, are closely related to the distribution of population. It is not surprising, therefore, that over 60 percent of nitrogen oxide emissions occur in urban areas. In terms of sources, only about 45 percent of the nitrogen oxide emissions from motor vehicles are in urban areas because, unlike carbon monoxide and hydrocarbons, they are relatively unaffected by travel speed.

### Carbon Monoxide

Petroleum-powered motor vehicles are the largest single emitters of carbon monoxide. Although vehicular travel is evenly divided between urban and rural areas, traffic in urban areas, where slower driving speeds prevail, accounts for 70 percent of total motor vehicle carbon monoxide emissions.

---

*Chemical Engineer, Office of Research and Development, U.S. Environmental Protection Agency

Preceding page blank

## Table 1
## 1970 United States Air Pollutant Emission Contributions by Source Category
(Percent)

| | Sulfur Oxides | Nitrogen Oxides | Carbon Monoxide | Hydro-Carbons | Total Suspended Particulate | Fine Suspended Particulate |
|---|---|---|---|---|---|---|
| Stationary Combustion | 75 | 46 | 0 | 3 | 4 | 9 |
| Electric Utilities | 52 | 18 | | | 3 | 7 |
| Indusrtial Combustion | 15 | 14 | | 1 | 1 | 2 |
| Residential & Commercial | 7 | 3 | | | | |
| Pipeline Pumping | 1 | 11 | | 2 | | |
| Motor Vehicles | 7 | 52 | 65 | 51 | 1 | 6 |
| Incineration | 0 | 2 | 8 | 5 | 1 | 0 |
| Industrial Processes | 18 | 0 | 10 | 28 | 8 | 15 |
| Carbon Black Manufacturing | | | 3 | 4 | | |
| Chemical Manufacturing | | | | 1 | | |
| Grain Handling | | | | | 1 | |
| Iron and Steel | 1 | | 4 | 1 | 1 | 4 |
| Mineral Products | | | | | 5 | 8 |
| Nonferrous Metallurgy | 10 | | | | | 1 |
| Petroleum Refining | 5 | | 2 | 10 | | |
| Pulp & Paper Manufacturing | | | 1 | | 1 | 2 |
| Sulfuric Acid Manufacturing | 2 | | | | | |
| Solvent & Petroleum Product Evaporation | | | | 12 | | |
| Natural Sources | 0 | 0 | 17 | 13 | 86 | 70 |
| Forest Fires | | | 8 | 7 | 26 | 21 |
| Controlled Burning | | | 9 | 6 | 14 | 11 |
| Natural Dust | | | | | 46 | 38 |
| Million Tons/Year Total Emissions (all categories) | 34 | 23 | 149 | 35 | 136 | 5 |

### Hydrocarbons

Of the sources surveyed, in the absence of controls, motor vehicles would again be the largest single emitter category. There are fundamental difficulties involved, however, in simultaneously eliminating hydrocarbon, carbon monoxide and nitrogen oxide emissions from mobile sources. Specifically, the carbon monoxide and hydrocarbon emissions could be reduced by increasing the ratio of air to fuel to the point where more air is present than is required for complete combustion of the fuel to carbon dioxide and water. Maximum emissions of nitrogen oxides, however, would occur under such conditions. On the other hand, at very low air-fuel ratios, the nitrogen oxide emissions could be reduced but high concentrations of carbon monoxide and hydrocarbons would be produced. At the extremely high air-fuel ratios where all three emissions could theoretically be low, operating difficulties such as misfire and stalling would be encountered with most commercially available internal combustion engines. This would result in poor performance and high emissions of carbon monoxide and hydrocarbons.

### Total Suspended Particulates

As indicated in Table 1, the largest sources of particulates are natural dusts and forest fires. These sources account for an estimated 86 percent of the national atmospheric primary particulate loading on a mass basis, and are a substantial portion of background levels. Emissions from these sources are essentially beyond the scope of present air pollution control methodology. These natural emissions differ from that found in urban industrialized areas, in that a generally smaller percentage of particulates are in the respirable range (less than 3 microns). Similar fugitive or unconfined sources directly related to man's activity involve agriculture, mining, construction and transportation

The mass of emissions from most of the industrial sources may be highly controlled by the installation of presently available control equipment. In some cases, such as combustion of coal by electric utilities, over 90 percent collection efficiency is routinely achieved. Nevertheless, because of the large amounts of coal burned and its inherently large noncombustible particulate content, emissions from coal combustion remain a major portion of the total particulate emissions from man's activity.

## FINE PARTICULATES

The overall impact of fine particulate pollutants, broadly defined as particles less than 3 microns in size, on man's environment is not well known. However, evidence continues to point to the many negative aspects of fine particulate pollution. Research on health effects of air pollution indicates links between fine particulate pollution and health effects of varying severity. In addition to particle size, the chemical composition of particulates is an important factor in determining the effects of this type of air pollution. For example, the health hazard associated with inhaled airborne particles depends on: (1) the site of deposition in the respiratory tract, which is determined by particle sizle; and (2) the effect on biological tissues at the deposition site, which depends on chemical composition. The hazardous pollutant problem (e.g., mercury, lead, cadmium, vanadium) is also directly linked to fine particle pollution because many of the industrial processes that emit hazardous pollutants liberate them in the form of micron or submicron particles. On both a regional and global scale, fine particulate pollutants play a principal role in the transport through the air of a variety of hazardous substances. It is emphasized that the hazards posed to human health by trace amounts of toxic materials in the form of fine particulates can be disproportionate to the mass involved.

Fine particulate air pollutants may be classified into two major classes based on their origin. These are (1) primary fine particulates which are emitted or immediately condensed as fine particulates from a specific source, and (2) secondary fine particulates which are the products of atmospheric reactions to form particulate air pollutants.

## III. LOCAL EMISSION / AMBIENT AIR QUALITY RELATIONSHIPS

The major concern with source emissions is their impact on maximum ground level pollutant concentrations in the vicinity of the site. Several meteorological diffusion models exist which can be used to provide satisfactory engineering estimates of sulfur oxide and total suspended particulate concentration on any given time reference of concern. The results of these analyses can be used to define control and siting criteria based on the actual contribution of each source to ambient air quality.

Figure 1 summarizes this typical variation in contribution to ground level ambient sulfur oxide concentration for four emitter categories. For example, while utilities typically contribute over 50 percent of all sulfur oxide emission in a local region, well-known factors such as stack height may limit the corresponding impact on the ground level sulfur oxide concentrations within a few kilometers of the emitter to less than 30 percent of the total. Conversely, area and smaller industrial combustion sources typically contribute more heavily to the ground level ambient concentrations than would be inferred by their percentage of total regional emissions.

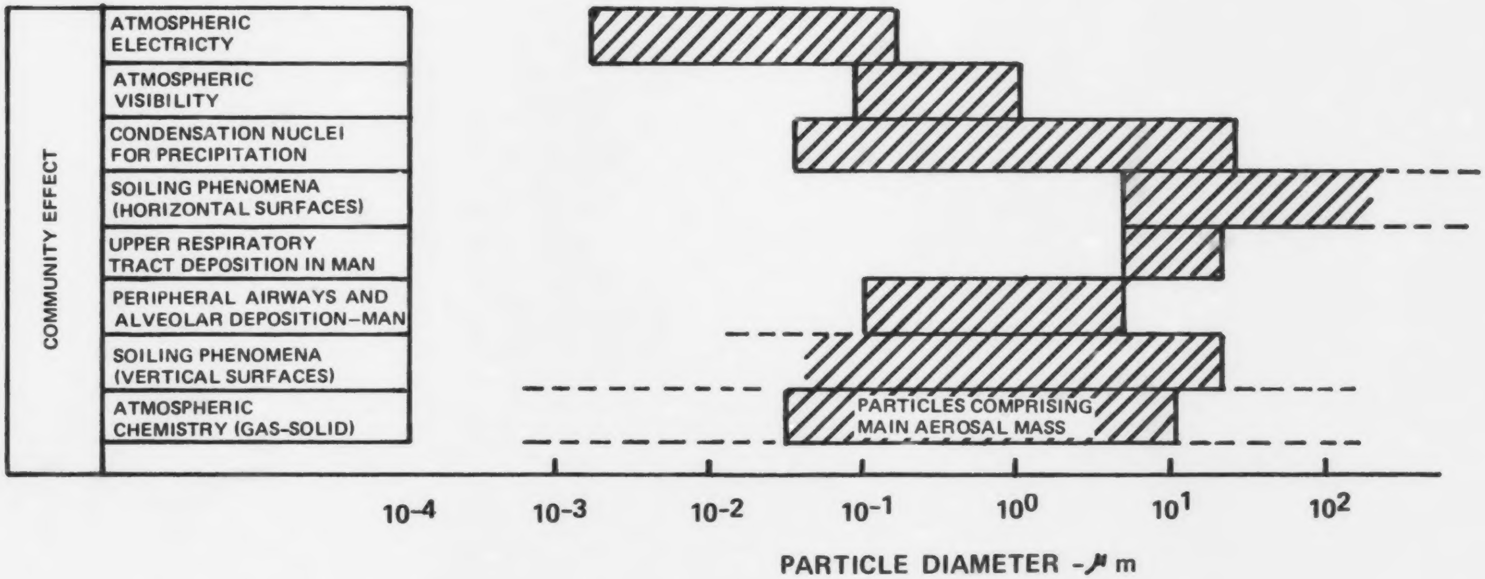
This ability to relate emissions to ambient air quality can be a very powerful and cost-effective tool for (1) determining the level of control required by source to achieve a desired ambient air quality goal, (2) allocating limited emission control capabilities to those sources having the greatest impact on ambient air quality and (3) defining new facility siting options compatible with achieving or maintaining ambient air quality goals.

Unfortunately, this technique has several limitations. First, it requires an accurate source-by-source emission inventory and regional meteorological data base. Second, considerable additional research is required before these meteorological diffusion models can accurately include the effect of atmospheric transformation primarily affecting hydrocarbons, oxides of nitrogen, oxidants and sulfates. Third, it may be concluded from the emission/ambient air quality relationships indicated in Figure 1 that peak ground level concentrations of sulfur oxides and total suspended particulates may be reduced to levels consistent with ambient air quality goals thorugh the relatively simple expedient of increasing effective stack height. Indeed, such considerations have formed the basis for "supplementary" control methods designed to achieve air quality goals in the absence of sufficient low sulfur fuel or control technology. This should be considered only as a near-term tactical expedient however. For example, our current knowledge of the significant health effect thresholds for sulfates and related fine suspended particulate matter infers the possible need for ground level ambient concentrations that may be eight to ten times lower than current practice. Under these conditions, the use of tall stacks is inadequate for achieving the necessary ground level ambient conditions and indeed results only in subjecting a larger ground area to excessive concentrations. This inadequacy is further accentuated in the long term as a result of a continued growth of society and its associated energy needs. Unless improved control techniques continue to be developed and applied, by the year 2000 it is forecast that $SO_x$ and fine suspended particulate emissions from industrial activities will increase fivefold. These emissions will not only dominate regional conditions, but could produce global concentrations exceeding acceptable health thresholds.

## IV. STATUS OF CONTROL METHODS

### Sulfur Oxide Control

Beginning in 1966, both industry and government started working seriously to develop processes capable of removing sulfur oxides from fossil fuel combustion flue gases. Because of the operating complexity and cost associated with these essentially chemical processing facilities, their practical application at the present state-of-the-art is limited to



SOURCE: CONTROL TECHNIQUES FOR NITROGEN OXIDES FROM STATIONARY SOURCES, AP-67, NATIONAL AIR POLLUTION CONTROL ADMINISTRATION, MARCH 1970.

FIGURE 1. EFFECTS OF PARTICULATE AIR POLLUTION IN THE COMMUNITY AS RELATED TO PARTICLE SIZE

large industrial and utility steam generators (greater than 100,000 lbs. steam/hr—equivalent to about 10 MW). As a result of this activity, about a dozen processes have been developed to the point where they are being studied on full-scale industrial facilities. Nine of these processes are variations of the slurry scrubbing concept using ash, lime, limestone, or magnesium oxide as the reactant. Development results to date generally indicate a sulfur dioxide removal capability of 80–90 percent. As of January 1974, 44 commercial and demonstration flue gas desulfurization installations were under construction or on order for use on existing and new electric utility steam generators in the United States. These installations encompass 18,000 MW of generating capacity at an average contract cost of $60 per KW of controlled capacity. This represents 9 percent of the coal and oil fired electrical generating capacity in the United States.

The most serious inherent limitation in the broad commercial application of these processes is disposal of the unmarketable waste gypsum by-product. Aside from the immediate spatial problem associated with the storage and handling of this waste product, disposal must be performed in a manner which does not cause water pollution.

The immediate question is the extent of commercial application which can be achieved by effluent treatment technology and the rate at which this application can be achieved. The extent of application is essentially a question of economics. Based on field surveys, it is theoretically possible to retrofit as much as 80 percent of all existing utility and large industrial coal and residual oil-fired combustion sources with effluent treatment processes for the removal of sulfur dioxides. In reality, however, considerations such as process packaging within available space, by-product transport and disposal, reactant availability, plant capacity use factor, gas, and size will limit the degree of application. For example, the range of annualized $SO_x$ control process capital and operating costs (incremental costs of control) required to satisfy 80 percent of existing utility and industrial combustion facilities extends from 10¢/M BTU of fuel consumed to over $1.00/M BTU. A practical upper bound on the ultimate extent of flue gas control application would be about 40 percent of large utility and industrial coal and residual oil capacity. This level of application would imply an incremental effluent control cost on the order of 60¢/M BTU.

It is clear that the combination of natural low sulfur fuels and flue gas desulfurization technology will not provide the quantity and size range of combustion capacity required to achieve existing national ambient air quality standards. Thus, a third category of control technology, fuel desulfurization, must also be considered. These processes for not only desulfurization, but total particulate cleanup of coal and residual oil are required to achieve ambient air quality standards. Their basic role is twofold:

(1) Provide control technology for combustion sources to which other alternatives are not technically or economically applicable;

(2) Provide synthetic, low pollution fuel augmenting the limited domestic supply of natural gas.

These approaches range from the application of currently commercialized mechanical coal cleaning technology, through chemical coal and oil desulfurization technology which maintains the clean fuel in its original physical state, to coal conversion technology for the synthesis of low- and high-BTU gas. The development cycle for these latter efforts, however, restricts commercial application until 1978–1985.

Given the current limitations in the availability of low sulfur fuel and control technology, the near term achievement of sulfur oxide ambient air quality standards can be achieved by application of the previously discussed ability to predict emission/ambient air quality on a source-by-source basis. This understanding can significantly reduce the number of sources requiring control. A long term benefit of such an approach is the ability to limit the very expensive implementation of sulfur oxide control technology to a few critical sources until a better appreciation of the technical implication of sulfate control can be developed. The effects of these considerations has been examined in three separate urban regions in which sufficient data were available.

For example, analysis of the 10 worst receptors in the Philadelphia region indicated an average annual sulfur oxide concentration ranging from 166 to 255 micrograms per meter$^3$. As shown in Table 2, if controls could be selectively applied in the Philadelphia region, it would be necessary to control, at a 75-percent effectiveness level, only 53 of the emitters encompassing 19 percent of the total regional emissions to achieve the existing sulfur oxide air quality standard.

Extrapolation of the trends indicated in the urban regions examined suggests that the existing sulfur oxide air quality standard could be achieved in the eastern United States by the selective control of a set of sources using 30 percent or less of all coal and residual oil consumed. This extrapolation has been supported by relating actual air quality measurements to source fuel modifications.

## Fine Particulate Control Technology

Four major categories of industrial particulate control equipment (electrostatic percipitators, fabric filters, scrubbers and inertial or centrifugal mechanical collectors) each having special characteristics that determine the suitability for a particular gas cleaning application, have been available for many

## Table 2

**Selective Reduction of $SO_x$ Emission to Achieve 80 $\mu g/m^3$—Philadelphia**

| Emitter Category | Present Emissions | | Number of Sources Requiring Control at: | | | |
|---|---|---|---|---|---|---|
| | | | 100% | | 75% | |
| | Sources | Tons/Day | Sources | Tons/Day | Sources | Tons/Day |
| Industrial Combustion | 238 | 555 | 3 | 172 | 10 | 293 |
| Industrial Processes | 165 | 300 | 5 | 41 | 13 | 124 |
| Utility Power | 21 | 1345 | 6 | 519 | 9 | 578 |
| Area Sources | 276 | 4378 | 11 | 148 | 21 | 222 |
| TOTALS | 700 | 6578 | 25 | 880 | 53 | 1217 |

years. This presently available control equipment has achieved only limited effectiveness in fine particulate collection capability and is generally not applicable because of material limitations in situations where co-contaminants must be dealt with.

Efficiency data indicate that the only conventional control devices capable of significant collection of fine particulates from industrial sources are high-efficiency electrostatic precipitators, venturi scrubbers and fabric filters. Existing data, as well as theory, on fine particulate collection efficiency (both on a mass and number basis) of conventional equipment indicate that inefficiency in the fine particulate range is an inherent characteristic of the equipment. Therefore, definite constraints will be imposed on the types of emission standards for fine particulates that can be implemented in the near term. These inherent limitations may require that process modifications or other alternatives be used to meet near-term regulations.

These are three basic avenues that might lead to improved control of fine particulates: (1) droplet condensation and impaction; (2) augmentation of commonly used collection mechanisms by additional forces such as thermal, diffusiophoretic and electrical which do not approach zero in the less than 3 micron size range; and (3) utilization of particle conditioning or agglomeration techniques producing particles compatible with conventional collection devices.

Substantial improvements in fine particulate emission could be achieved using currently available control equipment as indicated in Table 3. This table also indicates the relationship of emission level to plume opacity. Table 4 indicates that in most cases, installation of this best available technology will have relatively minor effects on overall production costs, generally amounting to less than one percent of the value of the product produced.

The collection of fiine particulate emitted directly from combustion or other high temperature processes

## Table 3

**Estimates of Fine Particle Emissions as a Fraction of Emission Standard**

| Source | Estimated Fine Particle Emissions | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Uncontrolled | | Present | | BICD[a] Standard | | 10% Opacity Standard | | 5% Opacity Standard | |
| | Lb/Hr | Ton/Yr | Lb/Hr | Ton/Yr | Lb/Hr | Ton/Yr | Lb/Hr | Ton/Yr | Lb/Hr | Ton/Yr |
| I. Coal combustion | 796,000 | 3,184,000 | 243,000 | 972,000 | 15,000 | 60,000 | 5,050 | 20,200 | 2,510 | 10,040 |
| II. Iron and steel | | | | | | | | | | |
| A. Sinter machines | 7,100 | 28,400 | 1,400 | 5,600 | 64 | 256 | 722 | 2,888 | 354 | 1,416 |
| B. Basic oxygen furnace | 400,000 | 1,600,000 | 43,600 | 174,400 | 2,560 | 10,240 | 1,540 | 6,160 | 800 | 3,200 |
| C. Electric arc furnace | 27,400 | 109,600 | 3,600 | 14,400 | 645 | 2,589 | 1,036 | 4,144 | 683 | 2,732 |
| III. Cement plants, rotary kilns | 243,000 | 972,000 | 44,300 | 177,200 | 2,000 | 8,000 | 25,200 | 100,800 | 12,200 | 48,800 |
| IV. Asphalt plants, dryers | 3,260,000 | 1,956,000 | 257,000 | 154,200 | 25,600 | 15,350 | 200,000 | 120,000 | 167,000 | 100,200 |
| V. Ferroalloy plants | | | | | | | | | | |
| A. Closed electric furnace | 92,000 | 368,000 | 11,300 | 45,200 | 11,300 | 45,200 | NC[b] | NC | NC | NC |
| B. Hooded open electric furnace | 129,500 | 518,000 | 84,100 | 336,400 | 15,800 | 63,200 | NC | NC | NC | NC |
| C. Unhooded open electric furnace | 72,000 | 288,000 | 60,800 | 243,200 | 28,800 | 115,200 | NC | NC | NC | NC |
| VI. Lime plants, rotary kilns | 75,800 | 303,200 | 22,000 | 88,000 | 1,200 | 4,800 | 10,500 | 42,000 | 4,800 | 19,200 |
| VII. Municipal incinerators | 15,000 | 37,500 | 13,100 | 32,800 | 490 | 1,225 | 3,400 | 8,500 | 1,900 | 4,750 |
| VIII. Iron foundry, cupola | 22,900 | 22,900 | 13,100 | 13,100 | 356 | 356 | 5,900 | 5,900 | 4,000 | 4,000 |

[a] BICD—Best installed controls (not necessarily the highest efficiency device available, but rather the best that is generally being installed at

[b] NC—Not calculated.

**Table 4**

**Summary of Fine Particulate Control Costs at BICD-Level Control (\$/Unit of Production)—at Capital Charge Rate of 0.20**

| Source | Annual Production Rate for Model Plant | Fine Particle Control Efficiency | Unit Cost or Value \$ | Control Cost/ Unit of Production |
|---|---|---|---|---|
| Coal-fired electric plant | 2.4 × $10^9$ kwh | 98.13 | 0.015 | 0.00012 |
| Municipal incinerator | 80,000 tons | 96.71 | 10.00 | 0.550 |
| Cement plant (rotary kiln) | 3.0 × $10^6$ bbls | 99.17 | 4.32 | 0.049 |
| Asphalt plant (rotary dryers) | 90,000 tons | 99.21 | 23.50 | 0.177 |
| Iron and steel | | | | |
| Basic oxygen furnace | 1.0 × $10^6$ tons | 99.70 | 187.00 | 0.259 |
| Lime plant (rotary klin) | 87,500 tons | 98.43 | 18.00 | 0.240 |

is only one portion of the fine particulate problem. Atmospheric fine particles which contain sulfate, nitrate and organic aerosols are largely formed during secondary atmospheric reactions involving the controlled pollutants subject to ambient air standards. Conversion of sulfur dioxide into sulfate is considered to be responsible for one-half to three-fourths of the atmospheric particles resulting from man's activities. The conversion of nitrogen oxide to nitrate particles accounts for 7 to 14 percent of the particulate burden produced by man, and the conversion of hydrocarbons has been estimated to account for much of the remainder. The relationships are understood in a general way, but they have not yet been placed on a quantitative basis. The nature of such formation may be through catalytic oxidation with heavy metal ions, ammonia-sulfate dioxide reaction in the presence of water, photochemical oxidation, and on heterogeneous gas reactions. As a result, it is anticipated that the control of this secondary fine particulate fraction may involve further control of direct links in the chain of precursors or modification to the physical and chemical parameters ultimately affecting formation. Table 5 summarizes current best judgment as to the short- and long-term thresholds for particulate sulfate as compared to total suspended particulate and sulfur dioxide. These data indicate that particulate sulfate has a threshold health effect an order of magnitude less than sulfur oxide and total suspended particulate.

The capabilities of current sulfur oxide control technologies are targeted at achieving moderate (30-70 percent) reductions in sulfur dioxide concentration in urban areas. Unless other controllable links in the precursor chain can be defined, the sulfate levels in Table 5 imply the need for a high efficiency control technology (greater than 95 percent) on large industrial sources and increased use of clean fuels in smaller combustion sources. Similarly, nitrate control may require improvements well beyond the current state-of-the-art in nitrogen oxide control technology.

Mobile sources contribute only about 3 percent of the total particulate. However, studies on the size distribution of automotive exhaust indicate that 90 percent or more of the mass emitted consists of particles smaller than 1 micron. Estimated emission rates for gasoline-powered vehicles burning fuel containing usual quantities of tetraethyl lead (2.5 gm.Pb/gal.) are on the order of 0.08 to 0.25 grams/mile. Lead free gasoline will be required in cars equipped with hydrocarbon control catalysts since lead emissions quickly deactivate catalysts. With these lead-free fuels and catalysts, particulate emissions will be in the 0.02 to 0.05 gram/mile range and may be dominated by sulfates and nitrates.

Emissions from aircraft turbines are significant in some areas. Estimates of Los Angeles turbine particulate emissions are one-third to one-fifth of automobile-generated particulates. Near approach and take-off corridors, aircraft particulates dominate the atmospheric aerosol. Particulate emissions may cover a wide range from 0.1 to 2 percent of the fuel burned, depending on engine type and operating mode.

## Nitrogen Oxide Control Technology

The control of nitrogen oxides, as indicated earlier, must consider the impact of both stationary and mobile combustion sources. The combustion of fossil fuels in air results in the formation of oxides of nitrogen both as a result of chemically bound nitrogen in the fuel and high temperature fixation of atmospheric nitrogen. In high temperature combustion processes, atmospheric nitrogen and oxygen play a dominant role in nitrogen oxide formation. Thermodynamic equilibrium highly favors the formation of NO over $NO_2$. Once formed at high temperatures, the NO can react with excess oxygen to yield $NO_2$. However, the residence time usually available in combustion equipment is too short for more than 5-10 percent of the NO to be oxidized to $NO_2$.

In principle, nitrogen oxide emissions from stationary sources can be controlled by stack gas cleaning methods or by combustion modification. Unfortunately, however, flue gas treatment methods are

## Table 5
## ($/Unit of Production)—at Capital Charge Rate of 0.20

(SHORT TERM)

| Effect | 24-Hour Threshold, $\mu g/m^3$ Sulfur Dioxide | Total Suspended Particulates | Particulate Sulfate* |
|---|---|---|---|
| Mortality Harvest | 300 TO 400 | 250 TO 300 | No data |
| Aggravation of Symptoms in Elderly | 365 | 80 TO 100 | 8 to 10 |
| Aggravation of Asthma | 180 TO 250 | 100 | 8 to 10 |
| Acute Irritation Symptoms | 340 | 170 | No data |
| Present Standard | 365 | 260 | No standard |

(LONG TERM)

| Effect | Annual Threshold, $\mu g/m^3$ Sulfur Dioxide | Total Suspended Particulates | Particulate Sulfate* |
|---|---|---|---|
| Decreased Lung Function of Children | 200 | 100 | 11 |
| Increased Acute Lower Respiratory Disease in Families | 90 TO 100 | 80 TO 100 | 9 |
| Increased Prevalence of Chronic Bronchitis | 95 | 100 | 14 |
| Present Standard | 80 | 75 (Geometric) | No standard |

* There is a growing body of evidence that adverse health effects are related to some form of sulfates. However, inadequate data on concentrations and compound-specific health effects make highly uncertain even a "best judgment" estimate of a threshold.

faced with major technological and economic obstacles and at present, no treatment process, either catalytic or scrubbing, appears viable. Thus, combustion modification is the only demonstrated method for the control of nitrogen oxide emissions from stationary fossil fuel combustion.

Control technology developments to date have demonstrated that flue gas recirculation is the most effective combustion control technique for nitrogen oxide emissions from nitrogen fixation and that staged combustion is the most effective technique for controlling fuel nitrogen conversion. The application of this technology has controlled the nitrogen oxide emissions from gas and oil fired utility boilers by more than 50 percent to a level of 150-250 ppm.

Reduction in emissions of up to 80 percent can be achieved by the combination of reducing excess air from 50 to 20 percent and recirculating 50 percent of volume of the flue gas. The mechanism for the reduction of nitrogen oxide emissions via flue gas recirculation is primarily a result of decreasing flame temperature and dilution of the available oxygen in the flame zone. Unfortunately, in large installations the amount of flue gas recirculation desired poses major redesign problems for existing gas handling equipment and ducting.

In limited short-term testing, the above combustion modifications, combined with burner designs which prevent burner flame interaction, have also resulted in nitrogen oxide reductions of up to 50 percent in coal fired utility boilers. Experimental burner designs have resulted in further reductions of up to 90 percent. These burner design modifications are extremely attractive because they are very inexpensive and have little or no operating cost penalty.

In smaller households, commercial and industrial combustion sources, although the same design and operating principles apply, very little experimental data are available on the control of nitrogen oxide emissions from such sources. On a more long-term basis, work has been initiated on advanced combustion modification concepts, such as catalytic burning of natural gas and coal-derived synthetic fuels at temperatures near 1400°C. Ultimately, utilization of such new concepts will allow essentially pollution free energy conversion from fossil fuels.

Low excess air firing of new and existing units costs about $0.82/KW depending upon unit size, age and design. This amounts, however, to less than one percent of the average $225/KW cost of a new coal fired generating plant.

## Mobile Source Control of Nitrogen Oxides, Hydrocarbons and Carbon Monoxide

The standards for these motor vehicle related pollutants have been exceeded in a number of major urban areas. In the United States, urban areas representing nearly 55 percent of the Nation's population exceed the ambient air quality standards for one or more of these pollutants. Table 6 indicates the general range of relative contributions of emission sources in our urban areas.

**Table 6**
**Mix of Emission Sources in the Urban Areas—1971**

| Pollutant | *Percent of Total Emissions* | | |
|---|---|---|---|
| | Automobile | Trucks, Buses & Motorcycles | Stationary Sources |
| CO | 77–87 | 8–10 | 3–15 |
| HC | 50–65 | 5–10 | 25–45 |
| $NO_x$ | 40–50 | 8–13 | 37–52 |

Because of the fundamental problems discussed earlier in simultaneously controlling nitrogen oxides, hydrocarbons and carbon monoxide from current designs of moter vehicle internal combustion engines, the strategy in the United States depends on control of stationary sources to achieve nitrogen oxide air quality standards. Mobile source controls on the other hand will be applied to achieve hydrocarbon (oxidants) and carbon monoxide air quality standards. These controls may take the form of (1) emission controls on new and in-use vehicles, (2) the reduction of vehicle miles traveled through the use of traffic controls, mass transit, parking taxes, etc., and (3) the development and implementation of new power systems having improved emission characteristics.

In addition, uncontrolled hydrocarbon vapor losses from service stations will make this source as important a contributor to hydrocarbon emissions as the vehicles they serve. Translated into grams/mile, the hydrocarbon emissions from service stations exceed the 1976 new car hydrocarbon standards. Vapor displacement control techniques currently being developed can partially reduce these emissions by over 75 percent by 1977.

### *1. Emission Controls*

The cost effectiveness of these control techniques in terms of pounds of pollutant controlled per dollar expended on control are described in Tables 7 and 8 for hydrocarbon and carbon monoxide, respectively. The data in Table 7 show that control techniques designed to reduce hydrocarbon emissions from pre-controlled vehicles and service stations are clearly the most cost-effective. However, these are not necessarily the most important in eliminating a regional air pollution problem. The figures in parentheses in Table 7 show the reductions in total hydrocarbon emissions from the implementation of each control technique.[1] These data show that the use of pre-controlled vehicle retrofits in 1977 (Air Bleed or Vacuum Spark Advance Disconnect) will reduce total hydrocarbon emissions in the air quality region by only approximately 1 percent. Therefore, the cost-effectiveness of pre-control vehicle retrofits is very high, but they achieve relatively little in the way of improvement in air quality. Conversely, measures such as inspection/maintenance and catalytic retrofits, which are applicable to controlled vehicles, have relatively low cost-effectiveness but higher air quality impact. Unfortunately, as indicated earlier these catalytic devices may have serious and unacceptable impact on sulfate and on fine particulates. Gasoline marketing controls prove to be both important and cost-effective in eliminating regional hydrocarbon emissions.

**Table 7**
**Cost Effectiveness of Hydrocarbon Control Techniques***

| | *Date of Program Implementation*** | |
|---|---|---|
| | *1977* | *1980* |
| Inspection/Maintenance (all cars) | 1.64 (3.6 %) | 0.54 (1.5%) |
| Catalyst (1972-1974) | 1.79 (3.77%) | 1.31 (3.0%) |
| Catalyst (1968-1974) | 1.90 (5.5 %) | 1.23 (3.6%) |
| Air Bleed (pre-1968) | 5.48 (1.0 %) | 5.20 (0.5%) |
| Vacuum Spark Advance Disconnect (pre-1968) | 6.91 (1.25%) | 8.36 (0.6%) |
| Gasoline Station Controls | 10.67 (7.1 %) | 11.56 (9.6%) |

*Cost-effectiveness increases with the size of the figure.
**The cost-effectiveness of the 1976 hydrocarbon standard is 3.7.

Finally, it should be noted that inspection/maintenance is a prerequisite for all retrofit measures owing to the need to keep the retrofit devices in good operating condition. Hence, it is not possible in practice to select an approach such as Vacuum Spark Advance Disconnect, for example, based on its high cost-effectiveness and implement it without inspection/maintenance.

The cost-effectiveness calculations for carbon monoxide controls follow a progression similar to that for hydrocarbon controls. Retrofit devices for pre-control vehicles again prove to be the most cost-effective but measures for controlled vehicles have the greatest impact on air quality. However, the carbon monoxide emission reductions (shown in parentheses) achieved by those control techniques are considerably higher than those achieved for hydrocarbon. Air bleed retrofits, for example, yield a 7.1 percent reduction in total carbon monoxide emissions in 1977, while air bleed only yielded a 1 percent reduction for hydrocarbons. The relative importance of carbon monoxide control techniques is generally higher because regional carbon monoxide problems are caused primarily by motor vehicles. Again, inspection/maintenance is a prerequisite for retrofit.

---

[1] Calculations based on a region with 30 percent stationary source hydrocarbon contribution such as Philadelphia, Baltimore, and Indianapolis.

**Table 8**
**Cost-Effectveness of Carbon Monoxide Techniques***

| | *Date of Program Implementation*** | |
|---|---|---|
| | *1977* | *1980* |
| Inspection/Maintenance | 17.5 ( 7.7%) | 8.8 ( 6.3%) |
| Catalyst (1972–1974) | 20.0 ( 8.4%) | 15.2 ( 9.2%) |
| Catalyst (1968–1974) | 24.7 (14.5%) | 15.5 (11.8%) |
| Vacuum Spark Advance Disconnect (pre-1968) | 31.8 ( 1.1%) | 40.9 ( 0.7%) |
| Air Bleed (pre-1968) | 194.0 ( 7.1%) | 192.7 ( 4.9%) |

*Based on a region with 5 percent stationary source contribution such as Seattle, Phoenix and Minneapolis.
**The cost-effectiveness of the 1976 new car carbon monoxide is 32.9.

### 2. *Vehicle Mileage Reduction*

Increasing public awareness of the adverse effects of the automobile on the urban environment, and legislation such as the Clean Air Act that has resulted from this awareness, make it clear that urban development policies that have encouraged and relied upon unrestricted use of the automobile must be changed. Controls must be placed on automobile use; transit must be found to prevent future urban growth from generating large volumes of traffic. The need to reduce urban area auto use is no longer at issue. The problem is how to do so without excessively restricting the mobility of urban area residents.

Among the possible approaches to the solution of this problem, increased transit usage and quality seem to offer the greatest potential for success over the long run. Other possible approaches include increased carpooling, reducing trip frequencies or trip lengths, and direct vehicular restraints (e.g., vehicle free zones).

Planning and implementing substantial transit improvements are likely to present severe problems in the period 1973–77 owing to the difficulties of designing suitable transit systems and acquiring the necessary vehicles. Existing transit does not have the capacity to achieve large reductions in auto use. Most urban transit systems operate at more than 75 percent of capacity during periods of peak work travel. EPA calculations indicate that with this level of capacity usage, the maximum reduction in auto use that can be achieved by existing transit fleets is about 5 percent. Achieving a 10 to 20 percent reduction in auto use could require expansions of current transit fleets of at least 50 percent and possibly over 300 percent. Threefold fleet expansions in many urban areas could exceed the short run production capacity of the bus manufacturing industry. In addition to bus production problems, there appear to be significant short-run planning problems. Most existing urban area transit plans are projected to achieve decreases in auto use of less than 10 percent. The most ambitious transit plan that has come to EPA's attention, that for the Washington Metro System, is projected to be capable of achieving a 20-percent reduction in auto use if it were fully implemented in 1976. In fact, the system will not be completed until 1983. These production and planning problems suggest that although the potential for reducing auto use through improved transit is large, it may be unrealistic to expect reductions greater than 10 to 20 percent by 1977.

The cost of bus transit depends on the detailed characteristics of the bus system, notably on vehicle occupancies. Transit buses cost roughly $1.00 per mile to operate compared with $0.07 per mile for cars. Hence, a transit system that carries roughly 40 riders per vehicle round trip will cost about the same as the auto. Higher occupancy systems might achieve net savings of $100 per rider per year. However, with low occupancies cost could reach $900 per commuter per year. There is clearly a potential for achieving substantial emission reductions at a net cost savings through increased use of bus transit. However, precise cost estimates will not be possible until detailed plans for emission-control oriented transit systems have been developed.

Car pooling programs appear capable of achieving net cost savings. A car pool program for the Washington, D.C. area based on a locator system and increased parking fees has been estimated to require an initial investment of $1.3 million and to have operating costs of $0.6 million year. If this system achieves a three percent increase in auto occupancies for peak period downtown work trips, the savings it achieves in auto operating costs will equal the annualized costs of the system.

### 3. *Mobile Source Control Effectiveness*

The projections of emissions reflect reductions in emissions due to in-use vehicle controls; Federal new vehicle emission standards for automobiles, trucks, and buses; and Federal new source performance standards for certain categories of stationary sources. They also reflect substantial reductions in emissions from new and existing stationary sources due to present and planned State and local regulations.

Figure 2 gives a comparison of hydrocarbon and carbon monoxide ambient air concentrations now and that expected in 1977 based on the cumulative control capability of (1) inspection and maintenance, (2) control devices on new and existing vehicles, (3) 20 percent reduction in auto use, and (4) applications of best available hydrocarbon control technology to stationary sources.

## V. CONCLUSIONS

The preceding discussion of methods for the control of air pollution leads to the following conclusions concerning the application of these methods to control strategies.

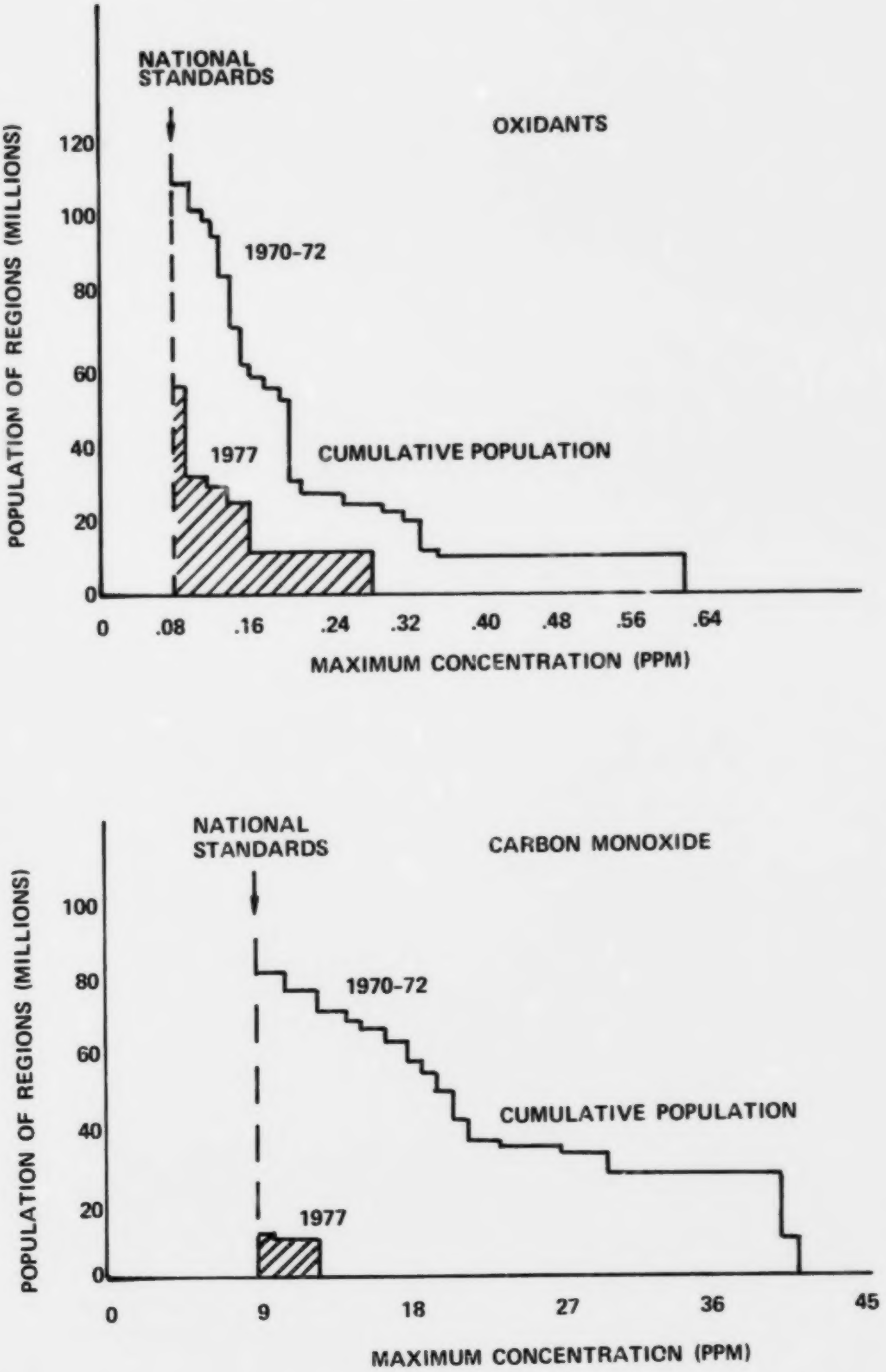


FIGURE 2. AMBIENT AIR CONCENTRATIONS

1. The achievement of sulfur oxide ambient air quality improvements is fundamentally dependent on the control of fossil fuel combustion, primarily coal and residual oil. In the near term (1975–1985), the availability of low sulfur content fuel and control technology limits reductions in urban sulfur oxide air quality to levels approximating current standards (60–80 micrograms per meter$^{3}$ on an annual basis). Even the achievement of this improvement must depend on the allocation of limited low sulfur fuel and controls to those sources imposing most heavily on ground level sulfur dioxide concentrations on a limited regional basis.

The potentially more serious problem of sulfate control may require widespread application of much more efficient sulfur oxide emission control methods than those currently being developed and commercially introduced. The apparently very low, safe ambient concentrations for this pollutant imply that plume dispersion considerations will not be appropriate for limiting source control requirements. As a result, until the interrelated questions concerning (a) what sulfate compounds impact seriously on health; (b) what are the atmospheric transport and transformation phenomena which produce these compounds; and (c) what precursors must be controlled to what levels are answered it may be very counterproductive to force the widespread installation of sulfur oxide control technology fundamentally inadequate to solve this problem at costs which often exceed 25 percent of the annualized cost of the controlled facility.

2. The widespread commercial installation and proper maintenance of relatively low cost existing particulate control technology can reduce the emission of fine particulate (less than 3 micron) from industrial processes and large combustion sources by 95 percent relative to current practice. This capability can be achieved at an increase of about 1 percent to 5 percent in the annualized cost of the controlled facility. The achievement of ambient air concentrations complying with probable future health standards for trace metals and carcinogens will, however, require further improvement in controls particularly for combustion sources, either through fuel pretreatment or flue gas conditioning. In either case, the technologies and energy requirements associated with further improvement imply a tenfold increase in control cost for every one percent further reduction in fine particulate emissions. Further, these improved capabilities are at early stages of development and will not be commercially available until at least 1980.

3. Current nitrogen oxide ambient air concentrations in the United States are generally consistent with current health based standards. As a result, emphasis is placed on a strategy of maintaining these levels. Combustion modification techniques capable of reducing nitrogen oxide emissions by at least 50 percent from all classes of stationary combustion sources appear to be either commercially available or in an advanced state of development. These techniques appear capable of at least maintaining current ambient air quality levels for the foreseeable future if adequate incentives for commercial implementation are applied. The costs of these combustion modification techniques are about one percent of the annualized cost of industrial and utility combustion facilities, but may be as much as 10 percent of the cost of smaller space heating equipment.

4. The control of hydrocarbons, oxidants and carbon monoxide is primarily a function of mobile source control strategies. If fully implemented in 1977, mobile source emission control techniques can produce a 20-percent reduction in hydrocarbon emissions and a 40-percent reduction in carbon monoxide emissions. This control capability will increase motor vehicle cost by about 5 percent. While these emission reductions and the corresponding urban ambient air quality improvements which result are significant, they are not adequate to universally achieve ambient air quality standards for oxidants and carbon monoxide. As a result, urban areas encompassing more than 40 million people in the United States will require at least 10 percent reductions in the vehicle miles travelled in addition to direct emission control techniques. Reductions of this magnitude will require expansion of urban mass transit capabilities extending past 1980 and incentives to shift urban commuter patterns to mass transit.

## BIBLIOGRAPHY

1. *Nationwide Inventory of Air Pollutant Emissions 1970*, U.S. Environmental Protection Agency, October 1972.
2. *Control Techniques for Particulate Air Pollutants AP-51*, National Air Pollution Control Administration, January 1969.
3. *Control Techniques for Carbon Monoxide, Nitrogen Oxide, and Hydrocarbon Emissions from Mobile Sources AP-66*, National Air Pollution Control Administration, March 1970.
4. *Control Techniques for Nitrogen Oxides from Stationary Sources AP-67*, National Air Pollution Control Administration, March 1970.
5. *Control Techniques for Hydrocarbon and Organic Solvent Emissions from Stationary Sources AP-68*, National Air Pollution Control Administration, March 1970.
6. *Abatement of Particulate Emissions from Stationary Sources COPAC-5*, National Academy of Engineering, 1972.
7. *Abatement of Nitrogen Oxide Emissions from Stationary Sources COPAC-4*, National Academy of Engineering, 1972
8. *Feasibility of Emission Standards Based on Particle Size*, U.S. Environmental Protection Agency Contract 68-01-0428, Midwest Research Institute, November 1973.
9. *A Survey of Emissions and Controls for Hazardous and Other Pollutants*, EPA-R4-73-021, U.S. Environmental Protection Agency, February 1973.

10. *Proceedings-Symposium on Control of Fine Particulate Emissions from Industrial Sources*, U.S.-U.S.S.R. Working Group Stationary Source Air Pollution Control Technology, January 1974.
11. *The Economics of Clean Air*, Report of the Administrator of the Environmental Protection Agency to the Congress of the United States, Document No. 92–6, March 1971.
12. *Environmental Considerations in Future Energy Growth, Vol. I, II, and IV*, U.S. Environmental Protection Agency, Office of Research and Development, April 1973.
13. *National Public Hearings on Power Plant Compliance with Sulfur Oxide Air Pollution Regulations*, U.S. Environmental Protection Agency, January 1974.
14 *Proceedings of Symposium on Multiple-Source Urban Diffusion Models AP–86*, U.S. Environmental Protection Agency, 1970.
15. *Some General Economic Considerations of Flue Gas Scrubbing for Utilities*, U.S. Environmental Protection Agency, Control Systems Laboratory, October 1972.
16. *Pollution Control and Energy Needs*, Advances in Chemistry Series 127, Chapter 5–T6, Effect of Desulfurization Methods on Ambient Air Quality, The American Chemical Society, September 1973.
17. *The Clean Air Act and Transportation Controls: An EPA White Paper*, U.S. Environmental Protection Agency, Office of Air and Water Programs, August 1973.

# Development of Maximum Permissible Environmental Loading (MPEL)—Water

Mark A. Pisano *

## I. INTRODUCTION

Ideally the desired level of water quality improvement and the date it will be achieved would be established by taking into consideration the costs of achieving various levels of improvement and selecting the level of control where the costs equate with the benefits. To this date the water pollution control program in the United States has not followed a strategy establishing a control level that in all cases would balance costs of control options with the benefits of reduced health and ecological damage.

In the 1948 Federal Water Pollution Control Act the United States established an enforcement conference procedure whereby control levels on individual dischargers were established by allocating pollutant load reductions in proportion to the reduction required to meet locally established water use (benefit) determinations. The process did not result in an adequate level of water quality improvement. Since that time there have been successive changes in national laws, most notably the Clean Water Restoration Act of 1966. This Act introduced the concept that the level of water quality to be achieved would be determined not by a local enforcement conference but rather national water quality standards were to be jointly established by States and the Federal Government. These efforts, however, also were not found satisfactory, owing partially to the complexity of the water pollution problem and the difficulty in implementing the regulatory machinery. Implementing a national pollution control program involves a large number of organizations and institutions, many of which have competing goals and objectives. To the extent that the implementation does not specify the decision rules and the target outcome of such a process, the resolution of these competing interests often fails. While national/state standards were established, the implementation of these standards was conducted on a case by case basis. The actual requirements levied against dischargers became unclear and the schedules for achieving the standards slipped.

In 1972, new legislation was enacted which attempted to solve the problems of previous efforts. This paper describes the method of establishing water quality targets and dates of their achievement and the design of the program for meeting these targets.

## II. REGULATORY STRATEGY

Previous papers have dealt with stress of pollutants on man and on ecosystems. These discussions contained, to the extent possible, an assessment of the risks and costs associated with stresses on the ecosystem, i.e., damage functions. From these functions target maximum permissible species stress could be established. Another set of papers concentrated on the control methods that are employed to achieve pollution control. These discussions also included the costs associated with these controls. Depending on the costs of these control methods and the time period required to implement them, a strategy specifying the level and time period of implementation of the Maximum Permissible Environmental Loading (MPEL) could be established.

The regulatory strategy in water is not based on such a balance where the costs of the level of control would equate with benefits. Rather the central thrust in the water program is that control measures themselves will establish the MPEL. Successive improvements in technology will be required over time, until a target of zero discharge is achieved. Thus, to a large extent, the "raison d'etre" of the program is a reliance on the advancing capability of technology. The program is based on an assumption that any discharge to water would upset the natural integrity of water, therefore any increase in the removal of pollutants is desirable.

The strategy does recognize that a uniform application of technology might not result in an appropriate level of controls to protect health and ecosystems in every case. Thus, the concept of a control level dependent on the instream criteria which support the intended use of each specific body of water

---

*Director, Water Planning Division, U.S. Environmental Protection Agency

Preceding page blank

is still maintained. If this water quality standard cannot be met by the requisite technology standard for the period, then a higher level of technology could be required. Also, if over time it is found that it is not technologically feasible and/or it is economically infeasible to move to zero discharge then the water quality standard can be used to determine a less strict level of control than zero discharge.

It should be noted, however, that successively higher levels of technological controls where feasible will be required even though the water quality standards might be achieved with less stringent controls; in which case the cost of control could be greater than the apparent benefits to society. This apparent inefficiency is permitted based on the assumption that if a pollutant can be removed at "reasonable" cost then it will be. Also, because of our historical inability to adequately equate cost of removing pollutants to the benefits achieved by cleaner water, a second best solution of at least requiring what technology can economically achieve regardless of the cost/benefit calculus has been an important part of the U.S. regulatory strategy.

There are two parts to establishing the MPEL in the U.S. The first is effluent guidelines which are like performance specifications as to what technology can do and when the technology can be implemented. The second is the water quality standard which serves as a guide or check point as to the adequacy of technology.

## Determining Effluent Guidelines (Table 1 provides a summary of these requirements.)

By 1977 all dischargers will be required as a minimum to install best level technology—secondary treatment for municipalities and best practical control technology for industries. The key concept is that by 1977, every facility of a given type and size must have its pollution release equivalent to all other facilities of the same type and size. To the extent possible, the objective is to develop guidelines that may be applied to individual dischargers without the necessity of treating each discharger as a unique situation. Because of the heterogenity of dischargers, industrial categories have been sub-categorized according to their age and size, raw materials, manufacturing processes, products produced, available treatment technology, energy requirements and costs.

Within each category the effluent amount is based on the known performance of the better performers having well designed and well operated treatment systems. A key consideration in the determination of the effluent limit for the category is a balancing of total costs of treatment, including capital costs and operation and maintenance, against the effluent reduction benefits achieved. The decision rule is to set the limit at the point where the marginal cost of removing additional amounts of pollutants significantly increases. For example, as shown in Figure 1, the incremental cost of removing additional amounts of BOD increases rapidly beyond the secondary treatment level for municipal discharges. This was a major reason for establishing the base level technology for municipal dischargers as secondary treatment and is expressed as shown in Table 2.

For industrial categories the removal levels and the parameters considered are not uniform between categories. These variations are shown in Table 3 which presents summary data for six effluent guidelines. Because of technological advances in some industries, the cost/effluent reduction benefit test results in a base removal level of zero discharge, e.g., insulation fiberglass.

**Table 1. Principal Statutory Considerations for Effluent Guidelines**

| Statutory Basis | General Description | Process Changes | Cost |
|---|---|---|---|
| **Best Practicable Control Technology Currently Available**<br>**[Existing Sources]** | **1. Achieve by 1977.**<br>**2. Generally average of best existing performance; high confidence in engineering viability.**<br>**3. Where treatment uniformly inadequate a higher degree of treatment may be required if practicable [compare existing treatment of similar wastes].** | **Normally does not emphasize in-process controls, except where presently commonly practiced.** | **Balancing of total cost of treatment against effluent reduction benefits.** |
| **Best Available Technology Economically Achievable**<br>**[Existing Sources]** | **1. Achieve by 1983.**<br>**2. Generally best existing performance but may include technology which is capable of being designed, though not yet in place; further development work could be required.** | **Emphasizes both in-process and end-of-process control.** | **Cost considered relative to broad test of reasonableness.** |
| **Standards of Performance Best Available Demonstrated Control Technology**<br>**[New Sources]** | **1. Achieved by sources for which "construction" commences after proposal of regulations.**<br>**2. Generally same considerations as for '83; more critical analysis of present availability.** | **Emphasizes process changes.** | **Cost considered relative to broad test of reasonableness.** |

**Figure 1. Typical Values, Biological Oxygen Demand (BOD)**

| | *Relative Cost* | *Carbonaceous* | *Removal* |
|---|---|---|---|
| Raw | 0.00 | 300 | 0% |
| Primary | 0.50 | 180 | 60% |
| Secondary (Hi-Rate) | 1.00 | 45 | 85% |
| Secondary (Conventional +) | 1.30 | 23 | 92% |
| Two Stage Nitrification | 1.75 | 23 | 98% |
| Advanced Waste Treatment + | 2.25+ | ? | 99% |

**Table 2. Base Level Technology Limits**

| Parameter | Measurement | Monthly Average | Monthly Average |
|---|---|---|---|
| Biochemical Oxygen Demand (5-day) | mg/liter | 30 | 45 |
| Suspended Solids | mg/liter | 30 | 45 |
| Fecal Coliform Bacteria | number/100 ml. | 200 | 400 |
| pH | units | within limits of 6.0–9.0 | |

By 1983 all industrial dischargers will be required to install best available technology (BAT). The level of control will be based on best performer of a category and will not include just end-of-pipe treatment, but process changes and product modifications as well. In some cases BAT may reflect prototype technology not yet demonstrated to be practicable. The objective is to develop technologies with higher removal rates, including some calling for "no discharge."

In setting the BAT limits, sub-categorization will also be employed. Within these categories costs will be considered relative to a broad test of reasonableness which includes the feasibility and cost of the technology, i.e., can the control be put in place at a cost that is not prohibitive. The economic test will not necessarily be the cost/amount of effluent reduced.

For municipal dischargers the 1983 requirement is not similar to the industrial standard. Instead of increasing the amount of reductions to be obtained by requiring more advance technologies, municipalities are required to consider alternative waste water management techniques and choose the most cost-effective and environmentally advantageous alternative. Treatment and reuse, and land treatment of wastes, both of which provide either zero discharge or high levels of treatment, must be among the alternatives considered. The environmental and cost-effective tests may demonstrate that a higher level of removal above base level technology or an alternative management technique may be required. The test is not uniform; rather a case by case assessment must be made.

## Setting Water Quality Standards

Water quality standards are established jointly by the States and Federal government. Each of these standards consists of a designated use criterion representing the acceptable limits of pollutants in receiving waters dedicated to that use. These criteria are based on the latest scientific knowledge of the effects of pollutants on human health and biological communities. With respect to uses, the U.S. has set as a goal that where attainable all waters will be swimmable and will provide for the protection and propagation of fish, shellfish and wildlife by 1983. The phrase "where attainable" recognizes that naturally occurring conditions, or uncontrollable non-point source pollution, could result in a failure to meet the 1983 goal. Table 4 provides an example of the 1983 use classification and criteria that are likely to be needed to support these uses. The criteria in this table are changing as new data are acquired. The process for setting these standards is to first establish the intended use of the water body and the intended biological activity desired. For these uses and biological protection the appropriate criteria (MPC) are defined, for example, temperature in a water body

**Table 3. BPT for Industry**

| *Industry* | *No. of Plants* | *Estimated Percentage Raw Waste Load Reduction Expected* | | *Other Pollutant Parameters Controlled* |
|---|---|---|---|---|
| | | *BOD* | *TSS* | |
| Cement | 166 | — | 84–100 | pH, Temperature |
| Meat Processing | 1200–4800 very small plants | 80–99 | 97– 99 | pH, Temperature, COD, TDS, Phosphorus, Nitrogen, Oil and Grease, Chlorides |
| Rubber | 92 | — | 80 | COD, Oil |
| Feedlots | 3500 large<br>1.6M–1.8M | 100 | 100 | COD, Chlorides, Nitrogen, Phosphorus, Fecal Coliform |
| Flat Glass | 47 | — | 75–100 | Oil, COD, pH, Phosphorus, TDS, Temperature |
| Insulation Fiberglass | 19 | 100 | 100 | Phenols, COD, TDS, Oil and Grease, pH, Temperature, Ammonia |

**Table 4. Example Water Quality Criteria Summary by Use Classification**

| Use Class. | Micro-biological | Dissolved Oxygen | Temp. | Hydrogen Ion | Dissolved Solids | Taste and Odor Producing Substances | Dissolved Gas | Color and Turbity Producing Substances |
|---|---|---|---|---|---|---|---|---|
| Class A Primary Contact Recreation (swimming, water skiing, etc.) | Shall not exceed a geometric mean of 200 fecal coliform per 100 ml. | Not less than 5 mg/l. Class B levels also apply. | 90°F Maximum. Class levels also apply. | Hydrogen ion concentrations expressed as pH shall be maintained between 6.5 and 8.3. | Shall not exceed 500 mg/l or one-third above that characteristic of natural condition (whichever is less). | None in amounts that will interfere with water contact use. | Class B levels apply. | Secchi disc visible at min. depth of 1 meter. |
| Class B Desirable Species of Aquatic Life and Secondary Contact Recreation (boating, fishing, etc.) | Shall not exceed a geometric mean of 10,000 total coliform or 2000 fecal coliform per 100 ml. (Fecal coliform counts are preferred.) | Not less than 5 mg/l (except for 4 mg/l for short periods of time within a 24-hour period). Not less than 6 mg/l in trout waters. Not less than 5 mg/l in marine waters. | *Cold Water*<br>(Trout) 50°F rise. Max. of 68°F.<br>*Warm Water*<br>(Bass etc.) 5°F rise in streams. 3°F rise in impoundments. Max. 90°F.<br>*Marine Water*<br>1½°F rise. | Hydrogen ion concentrations expressed as pH shall be maintained between 6.0 and 9.0. | Shall not exceed one-third above that characteristic of natural conditions. | Shall contain no substances which will render any undesirable tastes to fish flesh or in any other way make fish inedible. | *Cold Water*<br>Total dissolved gas pressure not to exceed 110 percent of existing atmospheric pressure. | *Cold Waters*<br>10 JU<br>*Warm Waters*<br>50 JU<br>*Marine Waters*<br>Secchi disc/visible at min. depth of 1 meter. |

will vary according to the level of biological protection. Finally, the flow regime defining the critical periods when standards must be met are specified, for example, 7-day, 10-year low flow in humid areas, 5-day, 5-year low flow in arid areas, etc.

## III. IMPLEMENTATION STRATEGY

The immediate task of the U.S. regulatory program is the classification of all hydrologic segments. Where the application of best practicable technology for industries and secondary treatment for municipal plants will result in meeting 1977 water quality standards, the segment will be categorized as an effluent guidelines limited segment. Where this control technology will be insufficient to achieve the necessary level of water quality, the segment will be classified as a water quality limited segment. Where there is significant doubt, a segment will be considered as a water quality limited segment, subject to a later reclassification.

About half of the segments are so heavily polluted the dischargers located on them will be required to go beyond 1977-level pollution controls to enable water quality standards to be met:

| | |
|---|---|
| Effluent guidelines limited segments | 1515 |
| Water quality limited segments | 1588 |
| Total segments identified by States | 3103 |

### Short Run Control Program

All dischargers must obtain permits which dictate the amount of discharge that is allowed to be discharged from an individual plant or municipality, and a schedule of remedial measures, including an enforceable sequence of actions or operations leading to compliance with the effluent limitation. Failure to comply with an interim or final requirement in the permit constitutes a violation of the permit, for which the permit holder could be fined up to $25,000 per day.

### Effluent Segments

The effluent conditions for dischargers in effluent segments are based on the effluent guidelines. A hypothetical example will be presented to illustrate the application of these guidelines. A proven and feasible technology for treatment of pickle-packer wastes is an "activated sludge treatment plant"; this is essentially just like a municipal sewage treatment plant, and is a good treatment technology for food-processing wastes. Furthermore, let us suppose that 20 percent of the dill pickle-packers already use this treatment method, and that the average performance of the best-run treatment facilities results in 1 pound of BOD (oxygen-demanding organic wastes), 2 pounds of suspended solids, and zero (no)-discharge of dill seeds or garlic cloves per 1,000 pecks of pickles packed. Thus our final effluent limitations might well be stated as shown in Table 5.

**Table 5. Hypothetical Effluent Limits**

INDUSTRY: Pickle-Packers
SUBCATEGORY: Dill Pickles

| | *Effluent Limitations* | |
|---|---|---|
| *Pollutant Parameter* | *Amount Permitted* | *Per Unit of Production* |
| **BOD** | 1 pound | Thousand pecks of pickles packed |
| **Suspended Solids** | 2 pounds | Thousand pecks of pickles packed |
| **Dill Seeds** | Zero Discharge | (No units needed) |
| **Garlic Seeds** | Zero Discharge | (No units needed) |

Given this sort of effluent limitation, we could easily calculate the pounds of BOD and suspended solids a dill pickle-packer of *any* size could be allowed to release and still be consistent with this standard. For example, a pickle-packer processing 5,000 pecks of dill pickles per 5-day week would receive a permit allowing release up to 5 pounds of BOD and 10 pounds of suspended solids per week, or 1 pound of BOD and 2 pounds of suspended solids per day, from the dill pickle part of his facility. The permit would also specify the final date when the packer would be required to achieve this level of discharge, and interim compliance milestones, e.g., begin construction.

## Water Quality Segment Analysis

Violations of standards in water quality segments could be caused either by point sources which must be subjected to controls beyond best practicable treatment (BPT)/secondary treatment or by nonpoint sources which must be controlled. The implementation program in the U.S. currently concentrates on point source discharges. Ideally the water quality segment analysis would determine the most effective allocation of waste leads between all point and nonpoint discharges. In the short run, however, if the contribution of the point source dischargers cannot be distinguished from the nonpoint source dischargers, then base level technology will be applied to the point source discharge.

The analysis would be completed for each parameter which is in violation of water quality standards. The analytical process for performing the load allocation for these parameters is outlined in Figure 2. Each source contributing that parameter to the segment should be identified and alternative remedial measures considered. The final treatment/control strategy for the segment should reflect a combination of control methods which will meet water quality standards for all water quality parameters.

As noted in Figure 2, modeling is generally the appropriate method of ascertaining total maximum daily loads and determining the effects of the proposed alternative abatement strategies.

Water quality analysis through modeling enables planners to predict water quality under adjusted conditions of flow, temperature and pollutant loads. Hence, it provides a basis for load allocation and effluent reduction determinations.

Water bodies may be categorized in one of three categories: flowing streams, estuaries, and lakes and impoundments. Dominant transport mechanisms differ in each category; hence, different modeling techniques are appropriate for different categories. Further, the degree of sophistication of technique within each water body category may also vary according to conditions in the water.

The water quality analysis and prediction developed by any model can only approximate the actual water quality which will occur under the various suggested hypotheses. These assumptions will never be entirely correct for the distinct water body being analyzed. Hence, remedial measures (effluent reductions) based on the model predictions will not result in water quality exactly as predicted by the model. Since unnecessarily stringent measures may result in costly overbuilding and inadequate measures may fail to protect the aquatic ecosystem and achieve established water quality goals, model selection must consider the degree of risk to be accepted.

Table 6 summarizes the criteria for the selection of an analytical technique. The table lists criteria for each of the four levels of complexity. For each level of complexity, the table presents the type of problems and water bodies for which the level is appropriate, the planning characteristics (complexity and risk) associated with that level and the time required for a study.

- Type A simplified analysis generally has relatively few waste sources and an assumption is made that there are no changes in waste loading and hydrology with time (steady state).
- Type B steady state linear kinetics has several sources which have overlapping stream effects, steady state conditions still hold, i.e., no change in waste loading and hydrology.
- Type C transient linear kinetics has multiple sources which have overlapping stream effects, waste loading and hydrology vary linearly with time, e.g., storm event diurnal variations and variable waste discharges. Type D time variable non-linear kinetics has multiple sources which have overlapping two-dimensional stream, lake, or estuary effects. The internal mixing process represents a major problem.

Once a model is selected a set of worst conditions is assumed and streamflows and temperatures are fixed to arrive at a waste load allocation. By successive approximations of waste inputs, the maximum allowable effluent levels for each of the dis-

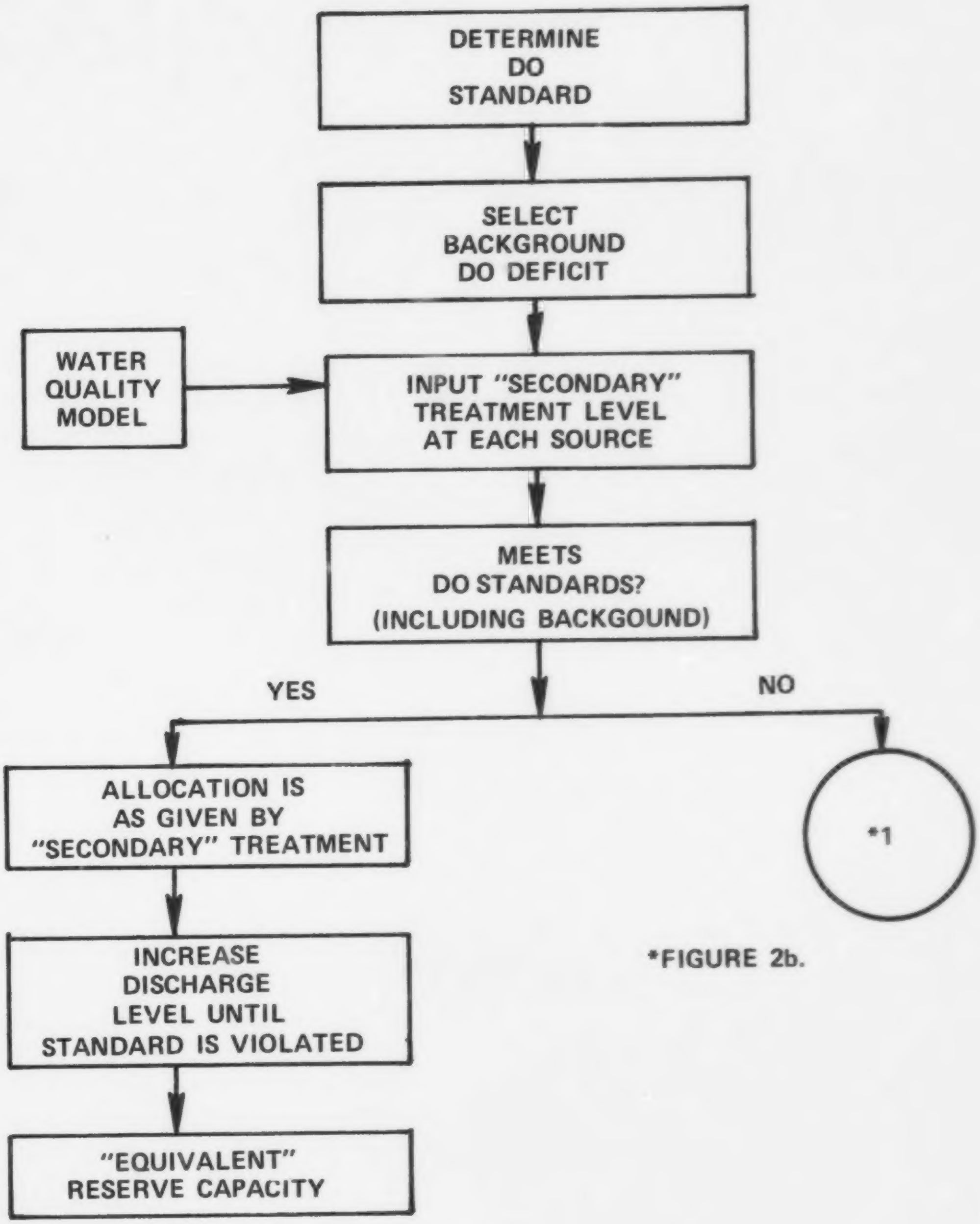

DETERMINE DO STANDARD
SELECT BACKGROUND DO DEFICIT
WATER QUALITY MODEL
INPUT "SECONDARY" TREATMENT LEVEL AT EACH SOURCE
MEETS DO STANDARDS? (INCLUDING BACKGOUND)
YES
NO
ALLOCATION IS AS GIVEN BY "SECONDARY" TREATMENT
INCREASE DISCHARGE LEVEL UNTIL STANDARD IS VIOLATED
"EQUIVALENT" RESERVE CAPACITY
*1
*FIGURE 2b.

FIGURE 2a. SUGGESTED ALLOCATION PROCEDURES (BOD – DO EXAMPLE)

chargers can be established to meet the water quality standard for a particular constituent. The goal may consider a reserve margin for accommodating future projected loads or uncontrolled natural background levels. The total allocation for a segment is the summation of all individual allocations of the component discharges.

In a system with a large number of pollutant sources the chance of each discharge being at its extreme allocated stress value simultaneously with all of the other sources and with the rare low design streamflow is statistically small. To issue permits based on the daily averages would be unnecessarily restrictive. It is reasonable therefore to define the permitted waste load for each discharger as a weekly average waste discharge.

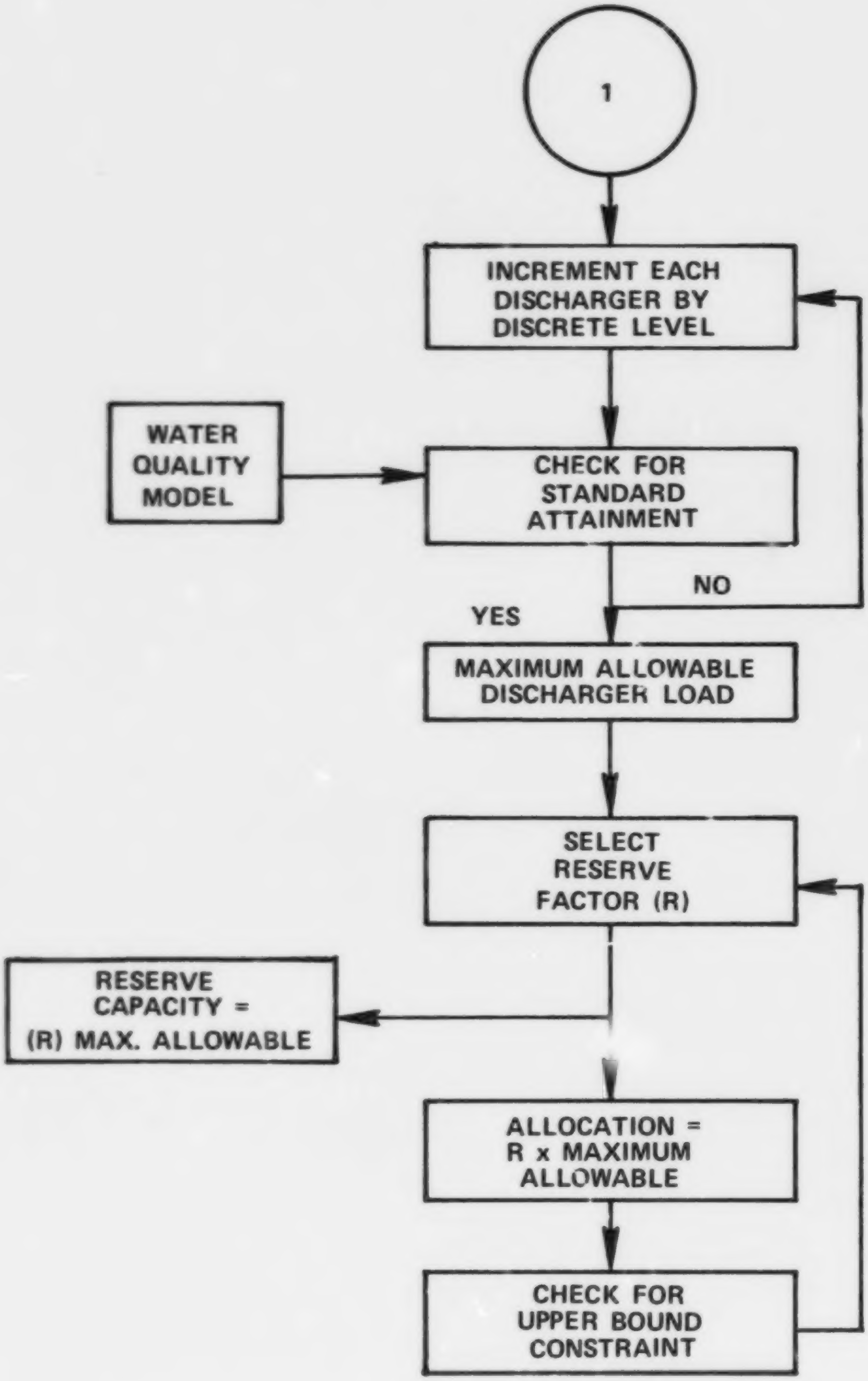

1
INCREMENT EACH DISCHARGER BY DISCRETE LEVEL
WATER QUALITY MODEL
CHECK FOR STANDARD ATTAINMENT
NO
YES
MAXIMUM ALLOWABLE DISCHARGER LOAD
SELECT RESERVE FACTOR (R)
RESERVE CAPACITY = (R) MAX. ALLOWABLE
ALLOCATION = R x MAXIMUM ALLOWABLE
CHECK FOR UPPER BOUND CONSTRAINT

FIGURE 2b. SUGGESTED ALLOCATION PROCEDURES, ctd.

**Table 6. Criteria for Selection of Techniques**

| Model Complexity | Water Quality Problems and Variable | Water Body | Planning Characteristics | Time Required for Study |
|---|---|---|---|---|
| Simplified analysis<br>(Type A) | D.O. (carbon and nitrogen) | One-dimensional streams and estuaries (completely mixed). | a. Low risk of capital and/or environmental quality degradation.<br>b. No alternate strategies and control options available. | Days to weeks. |
| Steady state linear kinetics<br>(Type B) | a. D.O (carbon and nitrogen), temperature and nonpoint source.<br>b. Anticipated or existing water quality problems. | One or two-dimensional streams, estuaries, rivers, lakes. | a. Low to moderate risk of capital and environmental quality degradation.<br>b. Alternative strategies and control options must be available. | 2 to 9 months. |
| Transient linear kinetics analysis<br>(Type C) | a. Time varying D.O., nonpoint source analysis, and temperature. Simple eutrophication analysis. Full storm water overflow analysis.<br>b. Water quality problems. | Rivers, lakes and estuaries. One or two-dimensional. | a. Moderate to high risk of capital and environmental quality degradation.<br>b. Alternative strategies and control options must be available. | 6 to 24 months. |
| Time variable non-linear kinetics analysis<br>(Type D*) | a. Detailed eutrophication analysis, etc.<br>b. Water quality problems.<br>c. High growth of area projected. | All bodies of water. | a. High risk of capital and/or environmental quality degradation.<br>b. Alternative strategies and control must be available. | 12 to 36 months. |

*Type D studies are to be performed only in very complex situations and should be regarded as research efforts.

Exceptions to this may be desirable when one or two dischargers dominate the combined load or when relatively large loads occur within short time periods (as opposed to reasonably steady day-to-day discharges). If there are only one or two dischargers in a segment, the probability of all peak loads occurring simultaneously is greatly increased. If the environmental risk is high, shock loads of some constituents should be restricted. For such cases, permits specified on a shorter term such as a daily mean rather than a weekly average would be prudent.

## Long Run Control Program

On some water quality segments the total loading discharged into the stream may be such that the best known technology may not be adequate to achieve the water quality standard when applied to existing point source discharges. On the other hand, the cause of the problem may be a non-point source problem. In many urban areas the control problem might be urban run-off or combined sanitary and storm wastes which bypass the treatment plant during periods of high flow. Construction of roads, housing and other facilities also creates problems of increased sediment run-off.

Where serious water quality control problems exist, areawide waste water planning and management agencies are being established. The agencies will develop control strategies that look not just at technology solutions, such as treatment facilities, but also will examine land use and land management alternatives. In many of these situations, altering or controlling future sources of pollution might be the only feasible solution. This might involve controls on the amount of growth that could occur in an area, or controlling where new facilities could locate. Or the solution might involve affecting the way in which land is used, e.g., the design of residential subdivisions.

In establishing the levels of controls for dischargers in water quality segments where higher levels than BAT are required, consideration of the social and economic costs and benefits of alternative control strategies is required. The control strategies could range from technological solutions to land use solutions. If there are no feasible solutions with costs equal to the benefits then the goal of swimmable waters and a balanced ecosystem may need to be relaxed.

## IV. CONCLUSION

Dr. Izrael pointed out that there are three methods of setting the MPEL:

- one that is ecologically desirable;
- one that reflects public health;
- one that we can technically achieve.

In the United States we are currently following a mixture of these approaches with an emphasis on doing what is technologically feasible, but not because a mixed strategy is necessarily better. Rather, our philosophy is that if it can be done, then do it, even if technology over-protects or under-protects public health and ecological objectives. As a scientific community we must find out what information the administrative process needs in order to fully implement a mixed strategy that balances all three of the limiting conditions for establishing the MPEL. To achieve this balance we need to improve our knowledge concerning concentrations of pollutants and their health and/or ecological effects. We also need to further our understanding of the frontiers of technology.

We also need to improve our ability to make better decisions on specific control requirements for individual dischargers. We need to improve our ability to make waste load allocations. It does not help the regulatory program if we have complete information on what the water quality criteria should be, but we cannot monitor or determine the control measures that are to be levied against point and non-point sources to achieve these criteria. A possible effort of future exchange could be improving the monitoring and analytical techniques used to relate water quality criteria to required control programs within a water body.

# The Effect of Polluting a Biogeocoenosis With Strontium 90 on a Population of Mammals and the Zoocoenosis

V. Ye. Sokolov, A. I. Il'yenko, D. A. Krivolutskiy *

Industrial pollution of the biosphere significantly affects the ecology of the organisms that populate the biogeocoenosis. This effect may be expressed by disruption of the biological processes taking place in the population, as well as by change in the relationships between organisms, and between organisms and the physicochemical factors of the habitat. Chemical pollution of the biogeocoenosis accumulated by its components can be considered to be a new abiotic factor for animals that acts on the population, community, and on the biogeocoenosis as a whole, along with the usual environmental factors of the habitat in which the organisms live.

The experimental introduction into the biogecoenosis of radioactive elements such as strontium 90, the chemical analog of Ca, for example, which is distinguished by the simplicity of determining the quantitative distribution of the isotope in a population, and among the components of the biogeocoenosis by radiometric and dosimetric methods, provides a way to establish the most vulnerable links in the animal ecology in the face of chemical pollution. The disruptions discovered in the ecology of populations and zoocoenosis in a biogeocoenosis polluted by strontium 90 can, possibly, serve as a model and be valid for many types of chemical pollution.

The question of the combined complex effect on animals of certain factors man has introduced into natural biogeocoenoses is being discussed widely today. We will devote special attention in our report to the effect that factors of the habitat which are normal for the natural populations have against a background of radioactive pollution of the biogeocoenosis.

The work was performed in two open-air cages, each with an area of one hectare. One of the cages was polluted with strontium 90 at a pollution density of from 1.8 to 3.4 millicuries/m$^2$. The other cage, which we shall conventionally call the "clean cage," had a background pollution consisting of artificial radioisotopes on a global level.

The cages were populated by two species of rodents: field (*Microtus agrestis*) and northern redbacked (*Clethrionomys rutilus*) voles. The beasts were trapped in the cages once a month, and were marked by amputating the toes in such a way that each zooid had its own individual number. Blood for study was drawn simultaneously, and part of the tail was severed for radiometric analysis.

Systematic observation of the tagged population yielded much data, and these are included in the report.

## The effect of polluting a biogeocoenosis with strontium 90 on the age structure of a population and the course of the biological cycles

The populations of the two species of voles kept under the conditions that prevailed in the open-air cage where the experimental section had pollution levels of 1.8–3.4 millicuries/m$^2$, showed differences in the degree of accumulation of strontium 90 -yttrium 90 in the skeleton (Il'yenko, 1967; Sokolov, Il'yenko, 1969). The concentration of isotopes in the boney tissue of the field (*Microtus agrestis*) and northern redbacked (*Clethrionomys rutilus*) voles was 0.11 ± 0.001, and 0.5 ± 0.001$\mu$curies/g, respectively, and the dose rate from the isotopes deposited in the skeleton was 0.77, and 0.35 rad/day for the bone marrow. The survival of the overwintered populations of these species of voles, that is for those who lived in the cage for more than a year, was not identical under these conditions (Figure 1). Evidently because of differences in the dose of irradiation from the isotopes deposited in the skeleton, the mortality was higher, and the overwintered population died off more rapidly, among the population of field voles, than was the case in the clean area. This pattern was less pronounced among the population of northern redbacked voles. The mortality among the population of field voles during the winter in the

---

*A. N. Severtsov Institute of Evolutionary Morphology and Animal Ecology of the Academy of Sciences of the USSR, Moscow

Preceding page blank

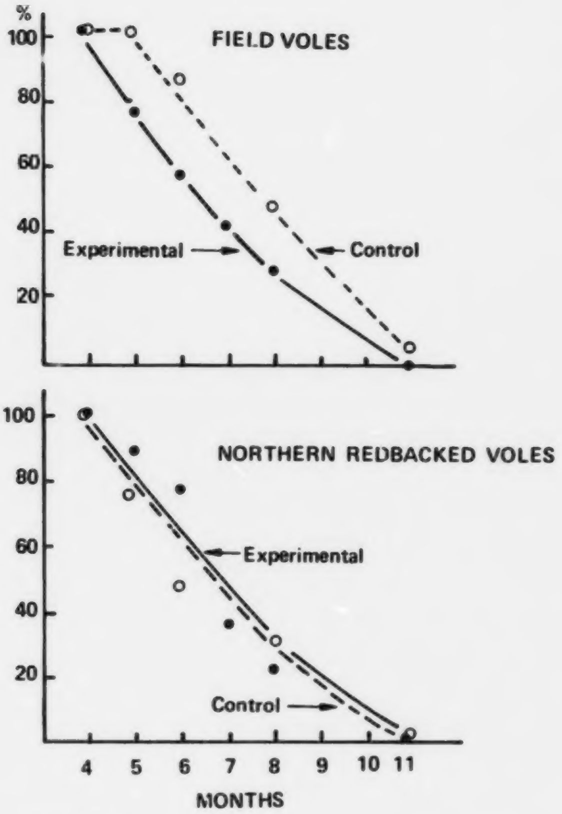


**FIGURE 1.**
**SURVIVAL OF OVERWINTERING POPULATION OF VOLES IN AN AREA POLLUTED BY STRONTIUM 90 AND A "CLEAN" AREA.**

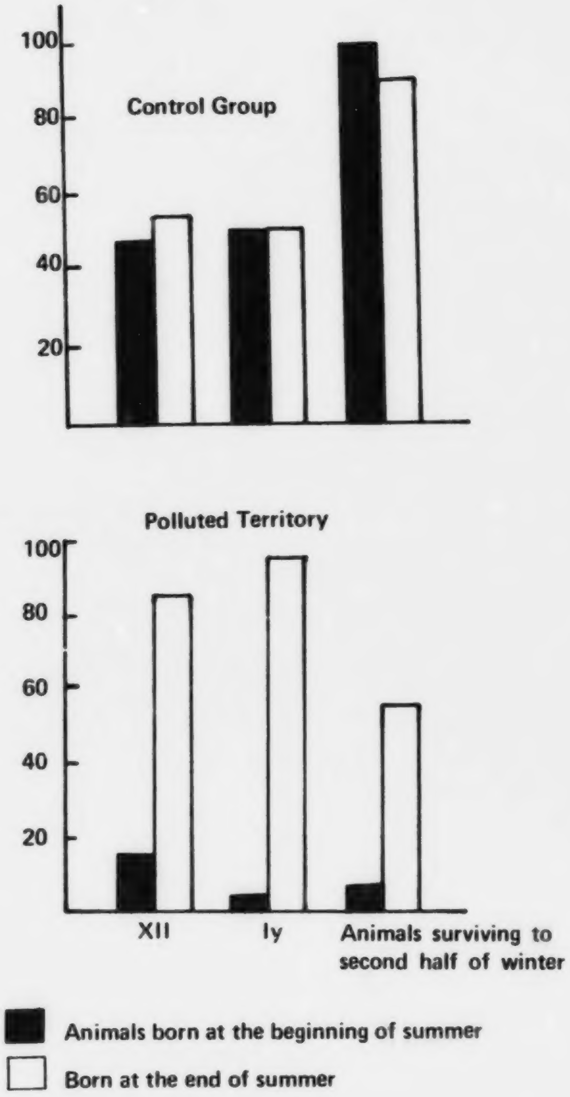


**FIGURE 2.**
**RATIO OF AGE GROUPS IN A POPULATION OF FIELD VOLES IN WINTER AND SPRING, AND THEIR SURVIVAL RATE IN THE POLLUTED TERRITORY AND IN THE CONTROL.**

polluted and clean cages was 95.2% and 50%, respectively. It is possible that differences in the radiosensitivities of these species also affected the mortality rate. Thus, the $LD_{50/30}$ for the northern redbacked voles is 960 rads, and is somewhere around 600 rads for field voles. The life span of field voles that survived the winter in the populations under study differed little, and those differences found were statistically unreliable. The mean life span for beasts in the polluted section, and in the control, was 11.6 ± 0.8, and 12.5 ± .08 months, respectively.

The populations of small mammals receive different doses of ionizing irradiation at different seasons of the year. The result is that the survival rate of the groups of beasts in the population in the polluted area is unequal (Figure 2) among the field voles born during different reproduction periods. Most of the beasts born at the beginning of the summer do not live until the spring of the next year. The number of animals born at the end of a reproduction period was 98% of the spring population. Only half of the animals born at the end of the summer survive the winter and live until the spring of the next year. Few deaths are recorded among the control population, however. The structure of the rodent population in the polluted territory changes substantially as a result.

Disruption in the reproduction rate was established during the study of reproduction patterns in the rodent population in the experimental sections polluted with strontium 90 at levels of from 140 to 1200 millicuries/$m^2$. The ovum mortality rate before implantation increased with increase in the strontium 90 – yttrium 90 concentration in the skeletons of long-tailed mice (*Apodemus sylvaticus*) (Figure 3). The sharp increase in the ovum mortality rate, greater than that in the control by a factor of 3, was observed when the concentration of strontium 90 – yttrium 90 in the animals' bone marrow was 0.003 $\mu$curie/g. The number of embryos in the females among the population of northern redbacked voles decreased with an increase in the concentration of isotopes in the skeleton (Figure 4). The study of the biological patterns found among isolated populations under open-cage confinement revealed that reproduction among the field vole population ended a month earlier than among the control population of this species confined under analogous conditions, but on unpolluted territory.

## The Effect of Climatic Factors

The action of ionizing radiation on a population is particularly strong during the extremal period of the existence of small mammals, when such action intensifies the action of all the unfavorable factors. This is why the mortality rate for the rodent population is particularly high in the winter time in an area polluted by strontium 90.

The rodent population does not renew itself in the winter time, something that is characteristic of other seasons of the year, and is the result of these animals' reproduction. The population dies off rapidly during this period, the result of the effects of a number of unfavorable climatic factors. The most critical times in the lives of the rodents are the periods of establishment of the permanent snow cover, and of the melting of the snow in the spring. Winter weather conditions are a lesser cause of death among small mammals and act differently on different species.

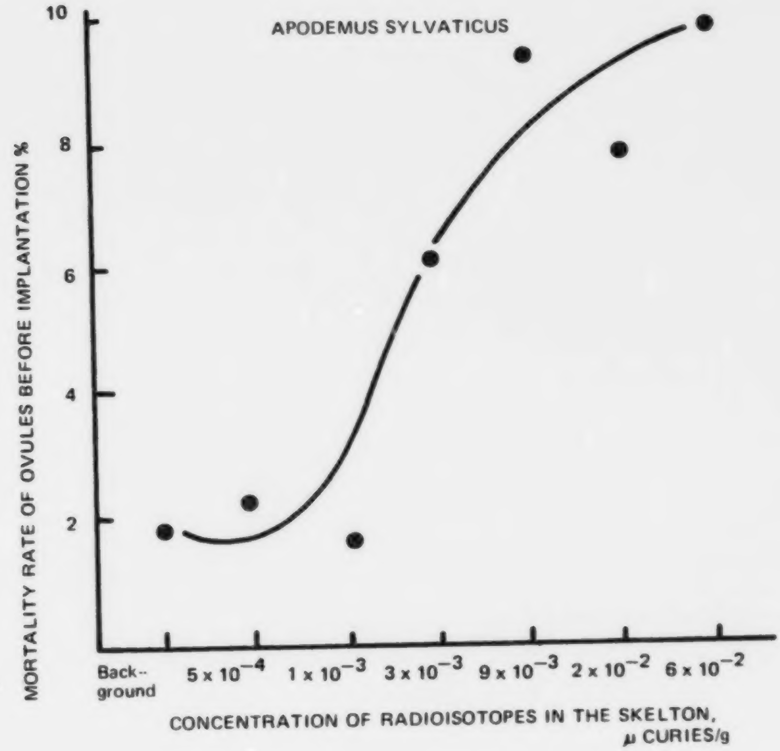


FIGURE 3. MORTALITY OF OVULES BEFORE IMPLANTATION IN A POPULATION OF LONG-TAILED MICE

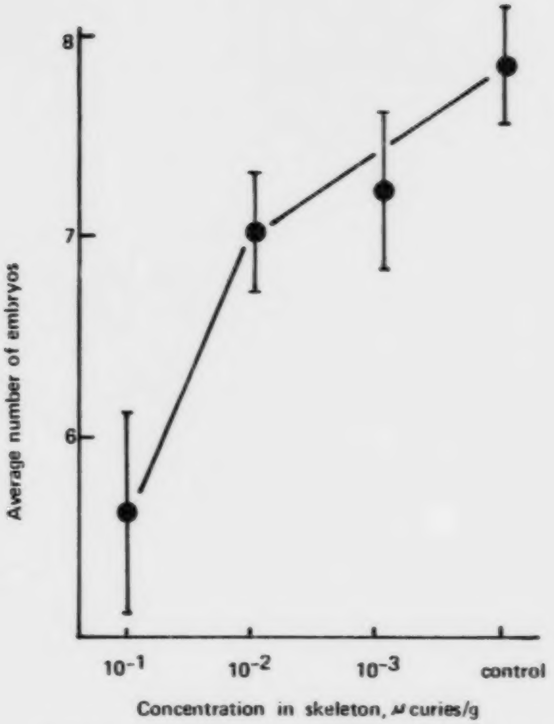


FIGURE 4. NUMBER OF EMBRYOS IN A POPULATION OF NORTHERN REDBACKED VOLES AS A FUNCTION OF THE CONCENTRATION OF STRONTIUM 90 - YTTRIUM 90 IN THE SKELETONS OF FEMALES

The data in Figure 5 were obtained from constant observation of individually tagged beasts kept in open-air cages. As may be seen, the mortality rate among the population of field voles on the territory polluted by strontium 90 was caused primarily by ionizing radiation during all seasons of the year, the effect of which is intensified by climatic factors. The autumn and spring mortality rates among the population of northern redbacked voles, a more radioresistant species already mentioned above that received smaller doses of ionizing irradiation, were determined by the effects of weather conditions unfavorable to the animals. Only in the wintertime was the more severe destructive effect of ionizing irradiation clear cut.

Thus, the effect of unfavorable habitat climatic factors on a population is intensified against a background of ionizing irradiation of the animals' bodies.

## The Effect of Biotic Factors

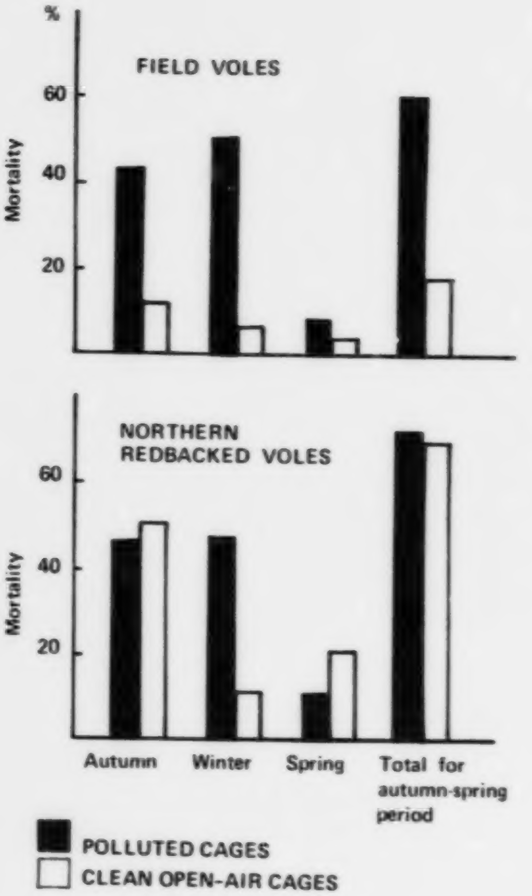


FIGURE 5. MORTALITY IN POPULATIONS OF VOLES IN A SECTION POLLUTED WITH STRONTIUM 90 IN WINTER.

The effect of the usual habitat ecological factors is intensified considerably when the population consists of herbivores, weakened by irradiation by ionizing radiation.

Study of the patterns of the use of a territory by a rodent population, using individual tagging of the animals and conducted by many researchers, has shown that the sections over which these animals forage will change in size, depending on the effect on the population of many of the usual habitat factors. Table 1 lists our data on the dimensions of the individual sections for adult male and female field voles during the most intensive reproduction period, midsummer. The population density was the same in the polluted and clean open-air cages.

The combined action on the females of ionizing radiation and the physiological load caused by pregnancy and feeding the young was expressed by a reduction in the area of the territory they used to a much greater extent than was the case for males affected just by ionizing radiation.

**Table 1**

**Change in the Area of Individual Sections in Populations of Field Voles Kept in Open-Air Cages in a Territory Polluted by Strontium 90, $M^2$ (Mature Animals)**

| | Polluted Section | | Clean Area | | Differences in area of sectors from mean dimensions in control, % |
|---|---|---|---|---|---|
| Sexual groups | Number of individual sections | Mean area of section | Number of individual sections | Mean area of section | |
| Males | 93 | 828±52 | 50 | 860±81 | 3.7 |
| Females | 134 | 473±15 | 41 | 696±43 | 32.0 |

Ionizing radiation causes a reduction in the natural resistance of the organisms to bacterial exo- and endotoxins and infections and parasitic diseases. The increase in the sensitivity of irradiated animals probably is caused by a reduction in their resistance to infections, and not by an increase in the virulence of the microorganisms in their bodies. A number of blood parasites were detected in the blood of the field voles we studied. Most were leukocytogregarines (*Leucocytogregarinae microti*). The number of beasts in whose blood endoparasites were detected was considerably greater in the population subjected to the constant action of pollution by strontium 90 during all seasons of the year than was the case among the control. Their summertime number exceeded the number of infested voles in the control cage by a factor of more than 6. None of the control animals were found to have blood parasites in the wintertime, generally speaking, whereas the number of infested animals was 16.5% in the experimental cage. Consequently, the effect on the rodent population of ionizing radiation from the strontium 90 deposited in the skeletons was expressed by a substantial increase in the number of beasts infested with blood parasites. Moreover, there was an increase in the number of blood parasites in the voles thus infected within the territory polluted by this isotope. Thus, 1 to 8 leukocytogregarines per 100 leukocytes were found in the blood of the control animals, whereas 40–60 were found in the blood of voles in the polluted area, and in excess of 80–100 specimens in individual cases.

The beta-radiation effect on the bone marrow from the radioisotopes deposited in the skeleton caused an increase in the number of leukocytes in the peripheral blood of the voles. The number of leukocytes was even higher among the beasts infected with blood parasites in the population of field voles living on the polluted territory. It seems obvious that the effect of blood parasites on the bodies of the voles is intensified considerably against a background of ionizing irradiation (Figure 6).

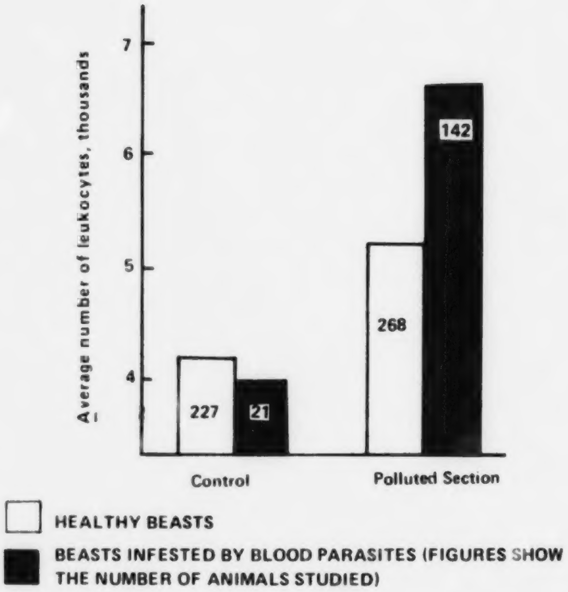


FIGURE 6. NUMBER OF LEUKOCYTES IN HEALTHY FIELD VOLES AND IN THOSE INFESTED BY BLOOD PARASITES.

The degree to which the various species of small mammals can become infested with parasitic mites is closely associated with the capacity of the beasts to put up an active defense against ectoparasites. Sick or weakened rodents with inhibited defensive capacities, usually are more heavily infested with ectoparasites than healthy animals.

According to our data (Il'yenko, 1971), the degree of infestation in the populations of four species of rodents, long-tailed and field (*A. agrarius*) mice, tundra voles (*M. oeconomus*), and northern red-backed voles, inhabitants of different biotopes, by ectoparasites, gamasid mites, had definite patterns.

Young animals that have lived briefly on polluted territory have a relatively small concentration of radioisotopes in their skeletons. The dose of ionizing radiation they receive from the isotopes deposited in the skeleton is still not enough to cause any changes in their bodies. Apparently this is why we detected no regular differences in the group of young animals in the degree of infestation by gamasid mites as a function of increase in the concentration of strontium 90 in the skeleton, and of the level of pollution by it of the biogeocoenosis. Mature animals that had lived for a long time in the polluted biogeocoenosis received a definite dose of ionizing radiation that corresponded to the quantity of isotope that had accumulated in the skeleton. The beta-radiation from the radioisotopes deposited in the skeleton can, in this case, cause disturbances in their bodies. The different groups of beasts among the adult rodents infested by gamasid mites are not identical. It was found that the number of mites on the infested animals, and the population infestation rate, increase with increase in the concentration of strontium 90 in the boney tissue. These patterns were found to be characteristic for populations of all four species of rodents.

Figure 7 shows the differences in the relative abundance of gamasid mites on the four species of mature rodents, which differences are associated with change in the dose-rate of ionizing radiation from the strontium 90 – yttrium 90 accumulated in the skeleton.

There is no doubt that irradiation of the bodies of small mammals by beta-radiation from the radioisotopes deposited in the skeleton weakens the animals, leads to inhibition of defense reflexes in protecting against ectoparasites, and causes an additional, greater effect by the latter on animals in the population.

## Changeability in Criteria in Animal Populations

As a rule, the effect of pollution of a locality by radioactive fission products on certain parameters of an animal population is studied by comparing the mean values of the criteria, or properties, of the populations being compared. At the same time, a number of researchers have shown that an equally important example of the biological effect of environmental factors on an animal population is the changeability, or the differing qualitative condition, of the population composition, reflecting the adaptability interrelationship between the organisms and living conditions and changes in these conditions. An evaluation of the extent of the effect of pollution on the population, expressed as the coefficient

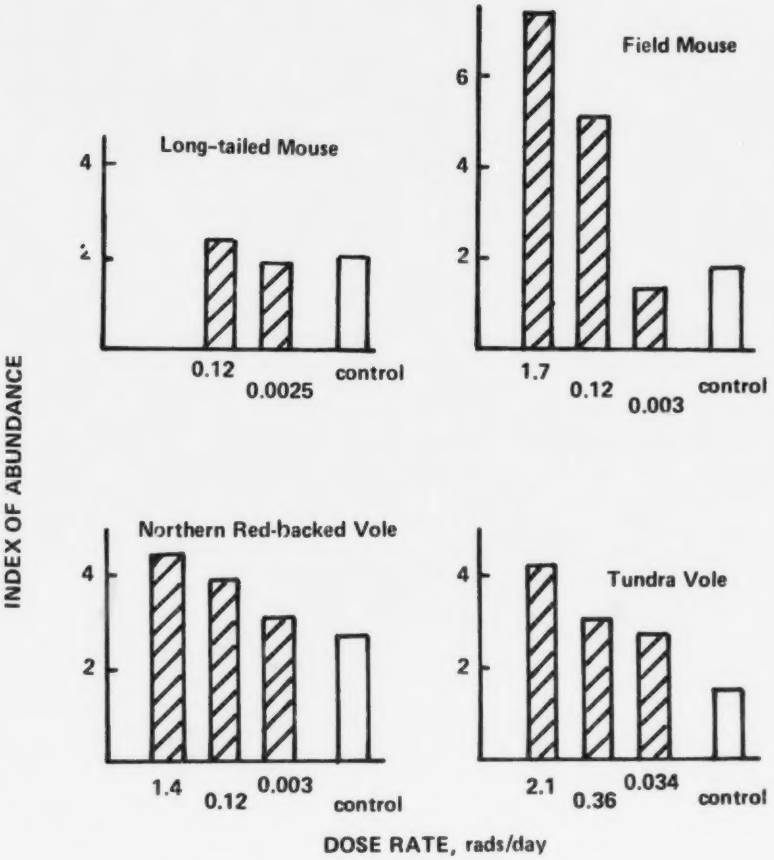


FIGURE 7. CHANGE IN THE INDEX OF ABUNDANCE OF MITES IN MATURE RODENTS AS A FUNCTION OF THE DOSE RATE OF BETA-RADIATION IN THE SKELTON

of variation, was found to be a more discriminating and reliable index than the mean value of the criterion (Il'yenko, 1968).

We studied the changeability in a number of population and morphological criteria in the populations of certain soil invertebrates and small mammals over an experimental territory polluted by strontium 90. An increase in the degree of changeability in criteria in the populations in the polluted locality was a general pattern for all the species studied.

Organisms incapable of active migrations for long distances were selected as the object for studying changeability in soil animals. It can be assumed that the populations in the polluted section consisted only of animals that had experienced the effect of ionizing radiation over the course of several generations, and which, moreover, had not been increased by immigrants from the surrounding territory. The observations were of two species of soil myriapods among the geophilids. A reliable increase in changeability in the number of pairs of legs was noted only for *Pahymerium ferrugineum* males. A tendency toward an increase in changeability was noted for the remaining groups, which makes it important to take into consideration the very rigid control conditions (Krivolutskiy, Smurov, Snetkov, 1972). All cases of deviation from the norm in the number of segments in the antennae (14 in this case) also were taken into consideration for all myriapods. The reliability of the difference in the sampled percentages of myriapods with these deviations was evaluated by the φ-Fisher method, and leaves no doubt as to the statistically reliable increase in the frequency with which deformities are found in the experimental population.

Figure 8 shows our data (Il'yenko, 1968; Sokolov, Il'yenko, 1969) on changeability in certain of the morphological criteria in populations of northern redbacked voles on an experimental territory polluted by strontium 90 at levels of from 0.6 to 3.4 millicuries/$m^2$. The coefficient of variation increases in the majority of the morphological criteria under these conditions of existence. This points to an increase in the degree of changeability in animals in a population inhabiting a biogeocoenosis polluted by strontium 90. Statistical processing of the data showed the high degree of reliability of their differences. A graphic representation of the distribution of criteria in the population on the polluted territory, and in the control, showed that the excess in the criteria distribution curve changes under the effect of ionizing radiation, and the amplitude of their changeability in the population subjected to chronic irradiation increases because of the appearance of a known number of animals with the criteria, the value of which exceeds the limits of changeability in the same indices for animals in the control population. Asymmetry of the criteria distribution curves will be seen in some cases, and this can be explained by the different effect of ionizing radiation on different animals in the population. The reliability of the differences in the distribution curves of the indices for the populations being compared was confirmed by checking in terms of the chi-square difference criterion. The numerical values of the coefficients of variation for different criteria in the population increase with increase in the concentration of strontium 90 in the skeleton of the northern redbacked voles. Consequently, the coefficient of variation can be a reliable index of the relationships between a radioactive environment and animal populations when studying the effect of radioactive pollution of a territory considered a factor in the organisms' habitat.

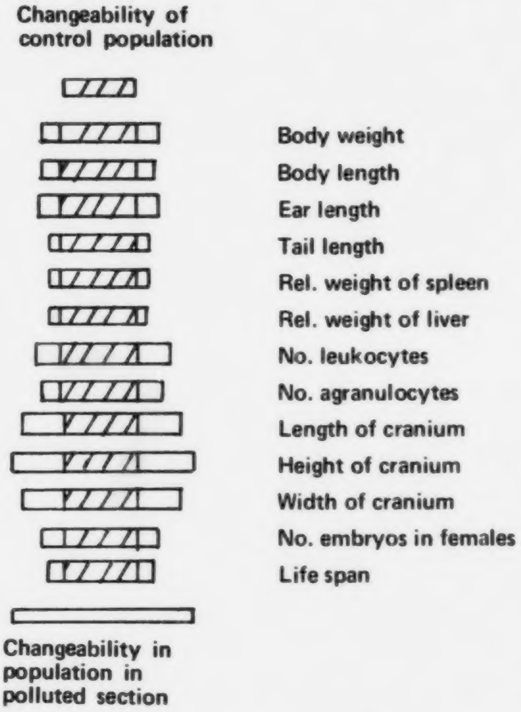


FIGURE 8. CHANGEABILITY IN MORPHOLOGICAL CRITERIA AND POPULATION INDICES IN A POPULATION OF NORTHERN REDBACKED VOLES IN A SECTION POLLUTED BY STRONTIUM 90.

The increase in the degree of changeability in interior (the relative weight of the spleen and liver, formed elements) and population indices (fertility indices, life span, and the like) serves as criteria for the degree to which pollution affects animals' bodies, and the population. An increase in the degree of changeability in exterior and craniological criteria, which are systematic criteria, may be indices of evolutionary transformations taking place in the popluation.

An increase in the changeability of morphological criteria in an animal population on a polluted territory shows that a change in the direction of natural selection can take place, because certain genotypes resistant to ionizing radiation will begin to predominate numerically in the population. Natural selection under these conditions inevitably leads, in time, to the occurrence of morphophysiological changes in the population, on the way to adaptation of populations to radioactive pollution of a locality.

## Effect on Zoocoenosis

Soil animals are one of the most convenient objects to use to study the effect of pollution on the animal population in a locality, because they are closely linked with the soil and vegetation cover, are relatively immobile, and comprise 70–80% of the total zoomass of land communities. Then too, study of the effect of pollution on soil animals is of definite interest. Soil animals deposit a goodly part of the active salts in their bodies and participate in soil-formation processes. The unfavorable effect of pollution on soil animals can be shown to be unfavorable for the entire biogeocoenosis.

The numbers of mesofauna were reduced by a factor of more than 2 in wooded sectors polluted by strontium 90 in a density of 1.8–3.4 millicuries/$m^2$. The differences were not the same throughout the warm season of the year. They reached a peak at the end of summer, the reproductive period for most of the groups of soil animals in the region under study (Gilyarov, Krivolutskiy, 1971). Not all groups of major invertebrates had their numbers reduced equally in the polluted section, something particularly characteristic of diplopods and earthworms. Saprophages, the consumers of the debris layer, the main component of the biogeocoenosis containing strontium 90, sometimes were virtually absent in the polluted sections. The number of predators decreased slightly on the whole, but sluggish groups, the centipedes in particular, completely disappeared from the population. The number of predatory beetles remained almost unchanged. This probably can be explained by the appearance of an additional source of food for the predatory beetles in the form of leaf-eating insects, a part of the life cycle, moving through the soil or plant litter. This additional source of food probably compensates somewhat for the direct action of radiation. Moreover, constant increases in the beetle populations, at the expense of surrounding sections, cannot be precluded. Noted in the section polluted by strontium 90 is extensive damage to birch leaves by leaf-eating insects, by gypsy moth, caterpillars, sawfly larvae, and others, in particular. In June 1969, 93.5% of the leaves here were damaged, as opposed to 2.5% in the control. Radioactive pollution of the environment also had an unfavorable effect on the development of the entire soil profile by soil animals. The vertical distribution of animals on a mean annual basis showed a general reduction in the soil population, particularly that in the deep levels, which cannot help but lead to weakening the soil formation rate.

Only the shelled mites of the microfauna in these same sections were studied in detail. Their numbers in the soils studied were 15–20 thousand specimens per square meter, and could not be differentiated reliably between the experimental and control sections. However, the variety of species of mites in the polluted territory was definitely decreased, particularly in the case of the species inhabiting the soil surface and the litter. Thus, what had appeared here was a tendency toward simplification of the coenosis in the polluted territory, even with respect to those animals whose overall abundance remained unchanged. This was particularly characteristic of the species inhabiting the litter and the upper soil levels (Krivolutskiy, Turchaninova, 1969). Observations of ants are of interest in studying the effects of pollution on complexes of land animals, because a considerable percentage of the biomass is made up of these animals, and their activities to a considerable extent regulate the number of leaf-eating insects. It is well known that ants react keenly to all forms of man's economic activities in the forest (improvement cutting, pasturage, burning of the litter, processing timber with agricultural chemicals). In the case of long-term radioactive pollution, we were unable to detect any diminution in the ant population in a quite large section in which the control and experimental sections were, in all other characteristics, practically identical, and which, prior to the introduction of the radioisotope, were a single section of the forest. It should be added that changes in the ant population caused by various forms of man's economic activities in the forest could be readily recorded in the research area by using the same procedure (Krivolutskiy, Baranov, 1972).

It is interesting that radioactive pollution of the environment has a strong, well-recorded effect, even on soil protozoa (Korganova, 1973), one of the most radioresistant organisms in nature. This effect can be explained only by the high level of radiation in the water films surrounding the soil particles attributable to the strontium 90 adsorbed by the soil.

## BIBLIOGRAPHY

Gilyarov, M. S.; Krivolutskiy, D. A., 1971, "Radioecological Research in Soil Zoology," *Zool. zh.* [Zoological Journal], 50, 3, 329–342.

Il'yenko, A. I., 1967, "The Biological Effect on a Population of Voles of Pollution of the Soil Surface By Strontium 90." In collection: *Ekologiya mlekopitayushchikh i ptits* [The Ecology of Mammals and Birds], "Nauka" Press, Moscow, 122–126.

Il'yenko, A. I., 1968, "The Effect of Pollution of a Section of a Locality by Strontium 90 on Changeability in a Population of Small Mammals," *Zool. zh.* [Zoological Journal], 47, 9, 90–97.

Il'yenko, A. I., 1971, "The Degree of Infestation of Populations of Small Mammals By Gamasid Mites in Sections Polluted by Strontium 90," *Zool. zh.* [Zoological Journal], 50, 2, 243–246.

Korganova, G. A., 1973, "The Effect of Experimental Pollution of the Soil on Soil Protozoa," *Zool. zh.* [Zoological Journal], 52, 6, 939–941.

Krivolutskiy, D. A.; Baranov, A. F., 1972, "The Effect of Radioactive Pollution of the Soil on an Ant Population," *Zool. zh.* [Zoological Journal], 51, 8, 1248–1251.

Krivolutskiy, D. A.; Turchaninova, A. A., 1969, "Some Changes in the Structure of the Animal Population of Forest Soils Under the Effect of Pollution by Strontium 90," In book: *Problemy pochvennoy zoologii* [Problems of Soil Zoology], "Nauka" Press, Moscow, 207–209.

Krivolutskiy, D. A.; Smurov, A. V.; Snetkov, M. A., 1972, "The Effect of Radioactive Pollution of the Soil by Strontium 90 on Changeability of Certain Organisms," *Zhurn. obshchey biol.* [Journal of General Biology], 33, 5, 587–591.

Sokolov, V. Ye.; Il'yenko, A. I., 1969, "Radioecology of Land Vertebrates," *Uspekhi sovrem. biol.* [Successes of Modern Biology], 67, 2, 235–255.

# Effects of Pollution on Species and Populations of Fish and Birds

Howard E. Johnson *

## I. INTRODUCTION

In very recent time man has attained the ability to cause major changes in his environment and that of other living organisms. As our technology and populations continue to grow human society has become increasingly aware of environmental limitations and our dependence on stable ecological systems.

The biota of an ecological system are the result of long evolutionary development where balance and equilibrium has been established between organisms and between these organisms and their environment. The different species within these biological communities have evolved together but each has developed rather specific physiological preferences and tolerances and they differ in many aspects of their life cycles from the reproductive stage to death. Frequently the species of greatest value to man are those least tolerant to environmental fluctuations or those highly restricted in their distribution in the environment. Even minor environmental change may be detrimental to their productivity or their very existence.

The strategy of pollution control to protect ecological systems differs from that where human health is a prime consideration. Rather than effects on the individual organism our primary concern is for protection of natural populations, functional biological communities or whole ecosystems. Relative to man, most animal populations have high reproductive rates and short life spans which produce a biological excess. Therefore, losses of individuals are not significant unless the fractional loss reduces potential harvest by man or is selective for a particular segment of the population. This generalization does not apply to species of birds or fish that have long life spans and low rates of population turnover.

Although the goal of pollution control is directed toward maintaining viable populations and the integrity of biological communities, our knowledge of factors regulating such complex systems is deficient. We can define the lethal threshold concentration of a pollutant for an individual species, but the pollutant concentration which will cause a major detrimental effect on a community or ecosystem is difficult to predict. Thus our immediate and most practical goal is to understand the effects of pollutants on individual species that are known to have major ecological significance or are of greatest economic benefit to man. If environmental requirements for these selected species are determined and appropriate controls implemented, we will have made a first major step in protection of the environment. Protection of these key species will also protect many other species that have similar or greater environmental tolerances.

Nevertheless the selection of beneficial or key species according to their economic or social value is disturbing to many ecologists. Some species not protected by the standards may be eventually eliminated from some environments with a resultant change in the biological community structure that is detrimental to even desirable species. As we reduce the diversity of an ecological system we often decrease its stability and internal regulation and require increased costs and energy in management. While this may be appropriate for some forms of agriculture or highly developed areas it is not feasible for all regions.

Thus our approach to ecological problems should necessarily be conservative. We should identify some major regions where human activity is highly restricted and retain these areas for the continued study of natural communities and to develop a reservoir of species that may be of major benefit to man in the future.

But social and economic demands for development of new regions require that pollution control standards are developed now according to our best scientific knowledge. Within any major region we can identify biological communities and species that are of major ecological or economic importance. Our research and control efforts should be developed to identify and protect these species. This includes determination of critical pollution concentrations as

---

*Associate Professor, Department of Fisheries and Wildlife, Michigan State University

Preceding page blank

well as the unique features of the environment that are required for their survival; for example, fish must have suitable conditions for spawning, migration, and food organisms.

## II. ESTIMATION OF MAXIMUM PERMISSIBLE LEVELS

The effects of most pollutants on an organism or a population can be described as a continuum extending from concentrations that are lethal after short term exposure to a level that has no adverse effects even after continuous exposure through an entire life cycle. Intermediate to these extremes, effects occur at different levels of biological organization that may or may not have ecological significance. Some changes merely reflect temporary or adaptive responses of individuals within their normal physiological range. Ideally, however, knowledge of the full range of dose-effect is important for decisions of cost-benefit analysis and to determine maximum permissible concentrations in the environment.

### Lethal Concentration of Pollutants

Because of their simplicity short term toxicity tests have been the method used most frequently to measure pollutant effects on fish and birds. In aquatic studies, fish are generally exposed to different concentrations of a toxicant for 4 days and the results expressed as the concentration that is lethal to 50% of the fish (96 hr LC–50). These data indicate the relative toxicity of compounds or the sensitivity of various species to a single compound, but the LC–50 data do not represent acceptable pollutant concentrations. It has been common practice to estimate maximum permissible levels of a particular pollutant by multiplying the LC–50 value by an application factor such as 0.1 or 0.05. The value of the application factor is arbitrary but it is usually based upon scientific judgment and field evidence (Sprague, 1971).

In the United Kingdom, investigations have shown an empirical relationship between the estimated short term toxicity (48 hr LC–50) of mixtures of common pollutants and the presence or absence of fish in certain rivers (Alabaster et al., 1972). In separate investigations with rainbow trout (*Salmo gairdneri*), Lloyd (1961), Herbert (1962), Herbert and Shurben (1964) and others have found the toxicity of some common pollutants is approximately additive. The proportional toxicity of each toxicant in the mixture is obtained by dividing its concentration in the water by the estimated 48 hr LC–50, e.g. $\frac{\text{Concentration of Toxicant A in Solution}}{\text{48 hr LC–50 of Toxicant A}}$ = the proportion of the lethal concentration of toxicant A and similarly for toxicants B, C and D in the mixture. The fractions for all toxicants in the mixture are summed to give the proportion of the 48 hr LC–50 for the mixture.

Using this technique in field studies, Herbert et al. (1965) found generally good agreement between predicted and observed toxicity in a number of fishless rivers in Great Britain. Alabaster et al. (1972) reported good fisheries generally existed in waters where the proportion of the 48 hr LC–50 was approximately 0.28 and marginal or no fisheries existed where the proportion was higher than 0.28. Brown (1968) reviewed the application of this method acknowledging the lack of scientific basis for additive toxicity of toxicants with different toxicological properties but suggested the various pollutants may each produce a non-specific but additive overall stress.

### Sub-lethal and Chronic Toxicity of Pollutants

There is general agreement among biologists for the need to consider the effects of pollutants at all life stages of an organism, particularly the effects on growth and reproduction. Mount and Stephan (1967) described a chronic toxicity test specifically designed to estimate acceptable toxicant concentrations for fish. The test provides for exposing groups of a single fish species to different concentrations of a toxicant through an entire life cycle in order to find the highest toxicant concentration that has no adverse effects on survival, growth and reproduction of the species. They suggest an application factor for the toxicant can be calculated by dividing the "no effect" concentration by the 96 hr LC–50. The application factor could then be multiplied times the 96 hr LC–50 for other species of fish in another water body to estimate acceptable concentrations.

The hypothesis for this approach is that the ratio of the "no effect" concentration and the 96 hr LC–50 of a particular compound does not vary greatly among different species even though each species may vary greatly in their sensitivity to the compound. Examples of acute toxicity values for several pesticides and heavy metals and the calculated application factors are presented in Table 1. In particular the data for malathion illustrate the close agreement of calculated application factors for different species that vary widely in their sensitivity to the insecticide. This tends to support the hypothesis that chronic toxicity data may have broad application to other species.

The validity of predicted toxicity from chronic toxicity tests in the laboratory requires verification in field studies. Eaton (1973) reported that preliminary results from studies of a natural stream intentionally polluted with copper seem to demonstrate the validity of predictions made from chronic toxicity tests with copper in the laboratory.

**Table 1.**
**Acute and Chronic Toxicity Values and Calculated Applications Factors for Some Heavy Metals and Pesticides**

| Chemical | Fish Species | 96-Hour $LC_{50}$ (mg/l) | Application Factor[a] |
|---|---|---|---|
| *Pesticides* | | | |
| Diazinon | Fathead Minnow | 6 | .0005 |
| Captan | Fathead Minnow | 0.065 | .10 |
| 2, 4–D Butoxy-ethanol ester | Fathead Minnow | 5.6 | .05 |
| Carbaryl | Fathead Minnow | 9 | .023 |
| Methoxychlor | Fathead Minnow | 0.0[illegible]5 | .107 |
| Malathion | Fathead Minnow | 10.5 | .02 |
| | Bluegill | 0.08 | .04 |
| | Brook Trout | 0.20 | .02 |
| Lindane | Fathead Minnow | 50 | .5 |
| | Brook Trout | 26 | .38 |
| *Heavy Metals* | | | |
| Chromium$^{+6}$ | Fathead Minnow | 33 | .03 |
| | Brook Trout | 50 | .01 |
| | Rainbow Trout | 69 | .003 |
| Copper | Fathead Minnow | 0.47 | .03 |
| | Bluegill | 1.1 | .02 |
| | Brook Trout | 0.1 | .09 |
| Cadmium | Fathead Minnow | 31 | .001 |
| | Bluegill | 20 | .0015 |
| | Green Sunfish | 20 | .0025 |
| Methylmercury | Fathead Minnow | 0.04 | .006 |
| | Brook Trout | 0.096 | .003 |

[a] The "application factor" is the ratio obtained by dividing experimentally determined maximum concentrations of a pesticide that have no detectable effect during chronic exposures by the 96-hour $LC_{50}$ for that species.

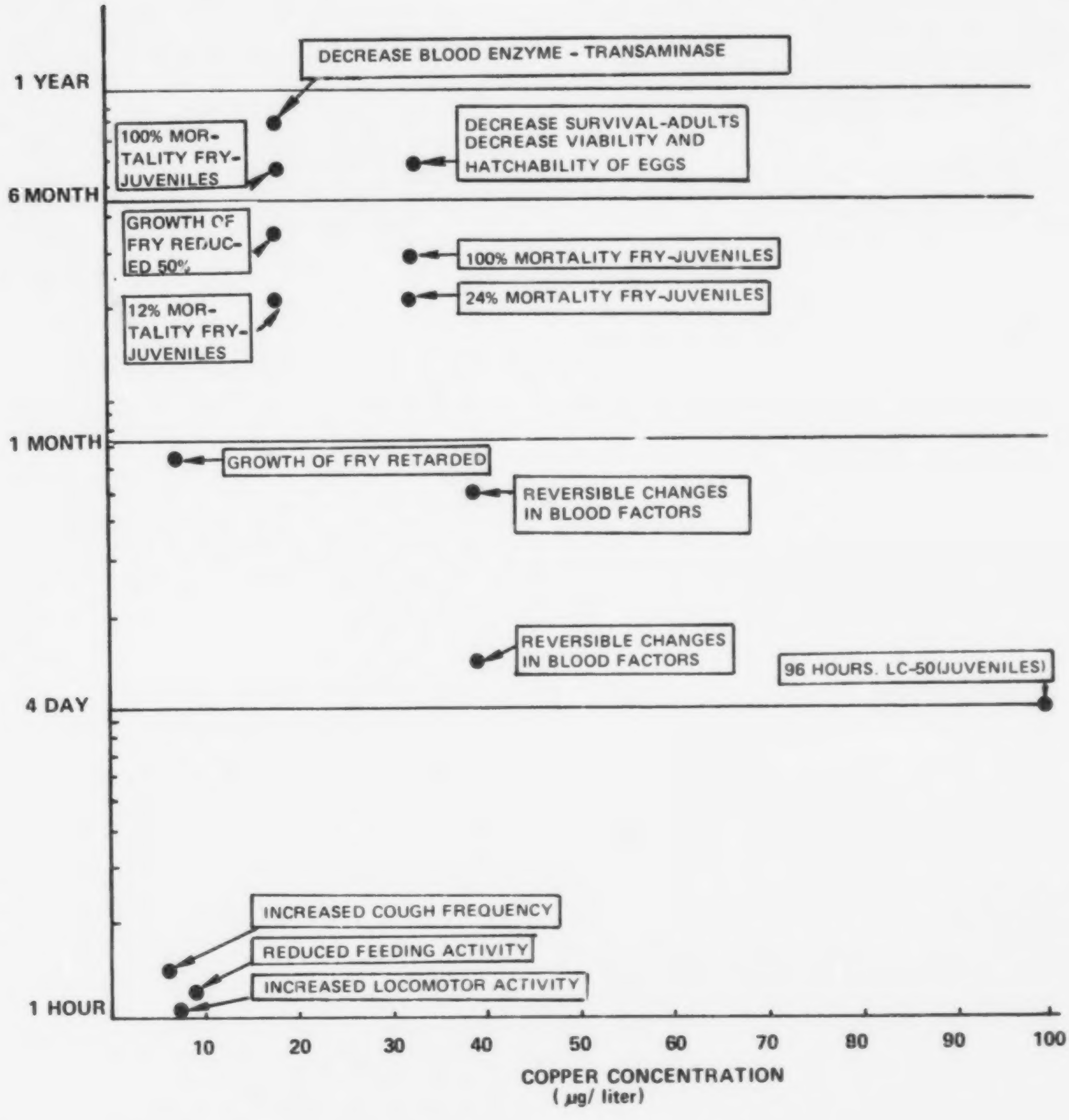
## Short Term Indicators of Chronic Toxicity

A major criticism of long term tests to define chronic toxicity or no effect levels is the long time and expense required for the test. It is obvious that such tests can be conducted with only a few of the major pollutants and species that are important. Although additional testing may show some data have broad application, other tests of shorter duration are necessary to predict long term effects. McKim et al. (1970) have recorded changes in several characteristics of brook trout blood after only six days exposure to copper concentrations that eventually affected reproduction, growth, and survival of fish after 8 months exposure (Figure 1). In subsequent tests, Drummond et al. (1973) recorded changes in behavioral activity of brook trout (coughing response, locomotor activity and feeding activity) that appear to be useful as short term indicators of copper concentrations that have no long term effects on the species. Cairns et al. (1973) have similarly reported that changes in locomotor activity and breathing movements of fish can be monitored as indicators of both lethal and sublethal concentrations of zinc in water.

Although chronic toxicity tests may reveal pollutant concentrations that affect reproduction, growth and survival of fish the practical significance of different degrees of effect must be determined; i.e., how important is 20% decrease in egg production or fry survival. Jensen (1971) used a mathematical model for yield of a brook trout population to examine the predicted effect of decreased egg production or increased mortality in "0" age groups, that occur due to a toxic pollutant. His results indicate that even a 5% increase in mortality of the "0" age group results in a substantial decrease in yield (Figure 2). More importantly his data indicate that an increased mortality in the "0" age group would not be detected as a significant change in the population for at least two or three years. Thus in natural fish populations, serious damage could occur due to effects on reproduction or survival in the early life stages long before field monitoring could identify the change as different from natural fluctuation in population density. These results emphasize again the importance of a conservative approach to setting standards for protection of fish and wildlife populations. Although threshold levels of effect can be identified, their values are usually related to only a few species and some allowance is probably necessary for protection of other more sensitive species including food organisms.

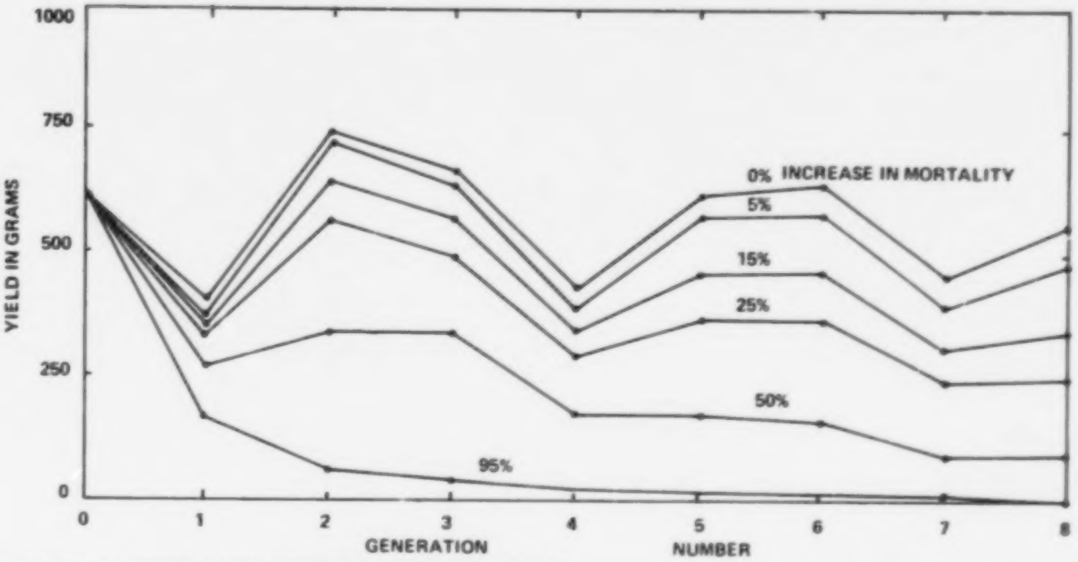
## III. EFFECTS ON BIOCIDES ON BIRD POPULATIONS

Bird populations can be affected by the introduction of biocides and other toxic materials to the natural environment. The species of birds and number of individuals in a particular region usually remain relatively steady year after year. A major objective of wildlife biology is to predict changes in the population as a result of observing changes in demographic parameters—fecundity and mortality—and to determine the cause of these changes. These parameters are usually measured by the number of eggs laid per clutch, frequency at which clutches are laid, survivorship of eggs and young to the age of first reproduction, and survival of adults throughout their lifetime (Ricklefs, 1973). These statistics of birth and death have been characterized for a hypothetical population by Ricklefs (1973) to demonstrate the relationship between recruitment, and adult mortality as a function of population density (Figure 3). In this example an annual mortality rate of 40 percent suggests that the population density will increase to a density of 125 individuals. As the annual mortality rate increases the density will decrease and the population will eventually become extinct if the annual adult mortality is over 80 percent. Although each bird species will have its own hypothetical model, these types of data are necessary in order to evaluate the effects of a toxic material that is released into the natural environment.



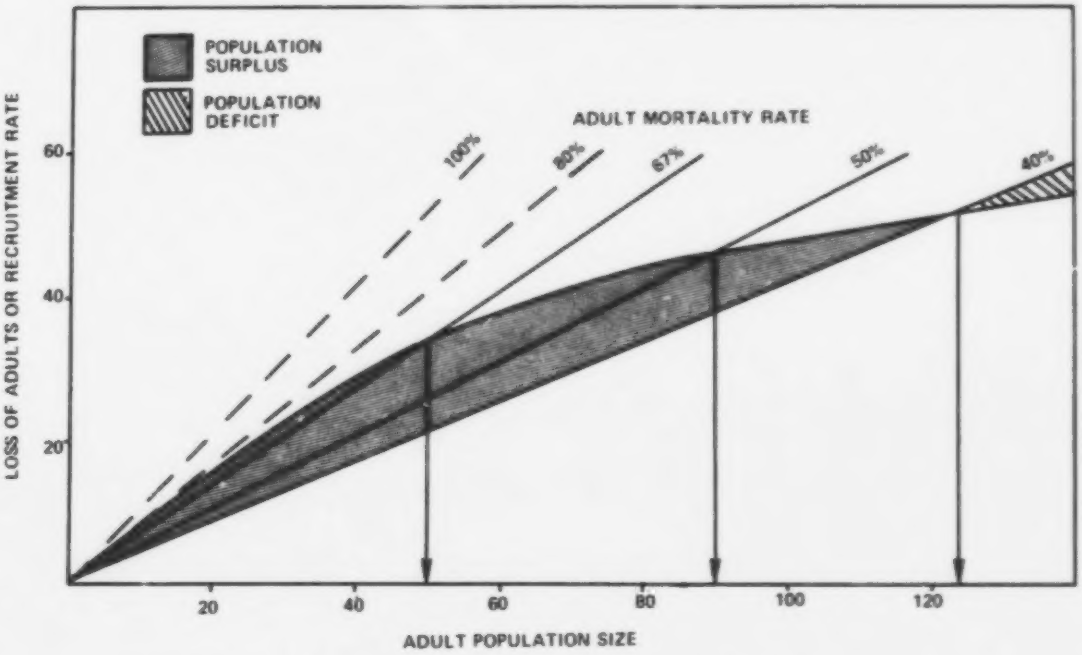
SOURCE: (DATA FROM McKIM et. al., 1970: McKIM AND BENOIT 1971, AND DRUMMOND et. al., 1973).

FIGURE 1. SOME EFFECTS OF DIFFERENT COPPER CONCENTRATIONS AND EXPOSURE TIMES ON BROOK TROUT *(SALVELINUS FONTINALIS).*



SOURCE: A.L. JENSEN, TRANS. AMER. FISH SOC. 100 (3): 456-459; 1971.

FIGURE 2. YEARLY YIELD FROM BROOK TROUT POPULATION WITH ADDITIONAL MORTALITY IMPOSED ON THE "O" AGE GROUP.



SOURCE: R.E. RICKLEFS, NATIONAL ACADEMY OF SCIENCES, WASHINGTON, D.C.; 1973

FIGURE 3. THE RELATIONSHIP BETWEEN ADULT MORTALITY AND RECRUITMENT AS A FUNCTION OF POPULATION DENSITY.

Most biological factors influencing recruitment and mortality in bird populations are related to the density of individuals. As density increases, mortality rates tend to increase and the rate of recruitment decreases. The reverse is true as density decreases. However, the effect of pollutants on bird populations is usually independent of density and mortality will continue despite a decrease in numbers. Survival is possible only if the population develops genetic resistance to the pollutant. Although there are no definite examples of genetic resistance to pollutants in birds, Ford and Prince (1973) have demonstrated a wide range of genetic variability among individual mallards (*Anas platyrhynchos*) to the effects of methylmercury chloride on the production of ducklings. Analysis for genetic variability has not been included in most evaluations of toxic materials, but should probably be included in future tests to improve predictions on the potential effects of pollutants on birds.

Different bird species vary widely in their response to specific pollutants, but usually species within a taxonomic order show a more uniform response than those between orders. Standard test species in the United States are mallards and ring-necked phesants *(Phasianus colchicus)* (Tucker and Crabtree, 1970). These species are common, widespread and important economically in this country.

Most pollutants are present in combination in the natural environment and present a potential for synergistic effects. Many of these effects will alter recruitment rates at lower concentrations than adult mortality rates. Thus data on reproduction is necessary to evaluate the environmental impact of most pollutants. A discussion of a specific pollutant, methylmercury, can best illustrate these considerations.

The wide use of mercury and mercury compounds in agriculture and industry has increased natural levels of mercury in certain regions. In Sweden starting about 1955, a decrease in the popoulation of seed-eating birds was attributed to the use of methylmercury compounds as seed-dressing agents. A similar instance of widespread mercury contamination in the Canadian prairie and depletion in the abundance of seed-eating birds, including pheasants, and their avian predators has been reported by Fimreite et al. (1970). A number of investigators have pointed out that a reduction in the reproductive potential could theoretically occur if pheasants consumed seed grain treated with alkylmercury compounds (Borg et al., 1969; Fimreite et al., 1970; Fimreite, 1971; Adams and Prince, 1972; Spann et al., 1972). Although only indirect evidence is available, this predicted decline occurred in Sweden, the Canadian prairies, and it may have been responsible for declines in pheasant populations in other parts of the range in North America. Removal of the alkylmercurials as seed dressing has apparently corrected these problems. A 1966 ban in Sweden resulted in an immediate drop in the merucury levels in the tissues of seed-eating birds and a corresponding increase in the number of birds (Borg et al., 1969; Johnels and Westermark, 1969; Jernelov, 1972).

When other seed-dressing agents such as sublethal daily doses of captan, dieldrin, diazinon, captan plus dieldrin, and captan plus diazinon are fed to laying pheasants reproductive declines are also observed (Stromborg and Prince, 1974). They found that decreased egg production was the most significant effect with dieldrin and captan acting additively, while captain and diazison showed significant synergism.

## IV. PERSISTENT POLLUTANTS

Some pollutants are of particular concern because of their wide dispersal in the environment and their tendency to accumulate in the food chain with potential adverse effects at high tropic levels including man. Most chemicals of this nature are synthetic organic compounds developed in the past few decades. The chlorinated hydrocarbon insecticides such as DDT, dieldrin, and heptachlor are examples but a number of other compounds widely used in industry also have properties of persistence, mobility and bioaccumulation. The polyclorinated biphenyls (PCB's) have been used in industry since about 1930 but their wide distribution and ecological hazard was not recognized until the late 1960's (Peakall and Lincer, 1970).

PCB's are used in a variety of products and manufacturing processes including synthetic resins, hydraulic fluids, lubricants and transformer fluids. Their loss to the environment has probably occurred through direct discharge of some fluids and oils directly to water, loss from dumps and incomplete burning in incinerators, and vaporization to the atmosphere. Through transport in the atmosphere and in rivers, PCB's have become distributed throughout the biosphere, and accumulated in various concentrations in invertebrates, fish and birds (Nisbet and Sarofim, 1972).

PCB's like other chlorinated hydrocarbons have a very low water solubility (less than 1 mg/liter). In surface waters PCB's have been reported at concentrations of parts per billion in a few highly industrialized rivers but most measurements indicate ambient levels of 10 to 120 parts per trillion. Acquatic invertebrates and fish can accumulate PCB to levels of $10^3$ to $10^5$ times those in the ambient water and sediments. Because these compounds are resistant to degradation they can be passed through the food chain to higher trophic levels. Additional accumulation occurs in birds and mammals associated with the aquatic environment to concentrations that are $10^7$ to $10^8$ times higher than ambient water concentrations (Nisbet and Sarofim, 1972; Stalling and Mayer, 1972).

The concentrations of PCB in fish and birds indicate high levels of contamination in industrial regions but low levels are widespread in other areas. Fish in some highly contaminated waters of the U.S. have concentrations ranging from 10–800 ppm but average concentrations in fish from various bays and estuaries or unpolluted rivers are 1 ppm or less (Panel on Hazardous Trace Substances, 1972).

The PCB's are generally considered to be only moderately toxic to animals with short term exposure. However, fish appear to obtain their greatest uptake directly from water and this appears to present a greater hazard than dietary sources. Due to their low solubility in water short-term tests of 96 hours or less do not adequately reflect their toxicity to fish or other aquatic organisms. The toxicity of the various PCB's varies between species but for 15 day exposure ranges between 50—30 μg/l PCB in water. With longer exposure the LC-50 values are decreased significantly (Stallings and Mayer, 1972). A pollutant which can be concentrated in organisms and is more toxic at longer exposure concentration presents an increased hazard to reproduction and growth. Some tests with fathead minnows (*Pimephales promelas*) indicate mortality of adults and reproductive failure after continuous exposure to concentrations of approximately 2 to 8 ppb PCB in water.

In the Great Lakes, the eggs of salmon and trout were reported to have concentrations as high as 12 to 17 ppm PCB (fresh weight) and may be partially responsible for the poor reproductive success of these species in Lake Michigan (Veith, 1969; Johnson and Ball, 1972). Johannson et al. (1970) reported high losses in prehatching salmon eggs that contain PCB residue of 14 to 34 ppm (lipid weight).

Other indirect effects of PCB on fish and bird populations have been reported. Hansen et al. (1971) found a high incidence of fungus infection in fish exposed to PCB in laboratory and Duke et al. (1970) found similar lesions in wild fish from an estuary grossly polluted with PCB. Ducks exposed to sublethal concentrations of PCB exhibited increased susceptibility to viral hepatitis (Friend and Trainer, 1970).

The difficulty of assessing the effect of a chemical pollutant on wild populations is compounded by the simultaneous occurrence of other pollutants or a change in some other environmental factor. High concentrations of PCB have been found in some wild bird populations that are predators on fish. Although these populations have experienced reproductive failure and population declines, the effect of PCB's is not clear because of the presence of DDT, mercury and other toxins. Where compounds affect reproduction, growth and behavior of organisms at concentrations below the acute lethal level, the effect may be expressed as a slow change in the size or relative abundance of a population. At present, levels of PCB in many parts of the environment occur at concentrations that exceed levels known to have adverse effects on organisms in the laboratory. Therefore, it is likely that adverse effects are occurring in populations or communities but it is very difficult to show this conclusively.

Many new chemicals which are potentially hazardous substances are introduced into commerce annually. Strategy for control of these pollutants must be directed at identification and testing for harmful environmental effects before they are used in large quantities or widely distributed in the environment. Many of our present techniques for testing the toxicity of new compounds are inadequate for evaluation of their environmental impact. Consideration must be given to testing in more sophisticated model ecosystems and biological communities where rates of transport, degradation, bioaccumulation and trophic transfer can be evaluated.

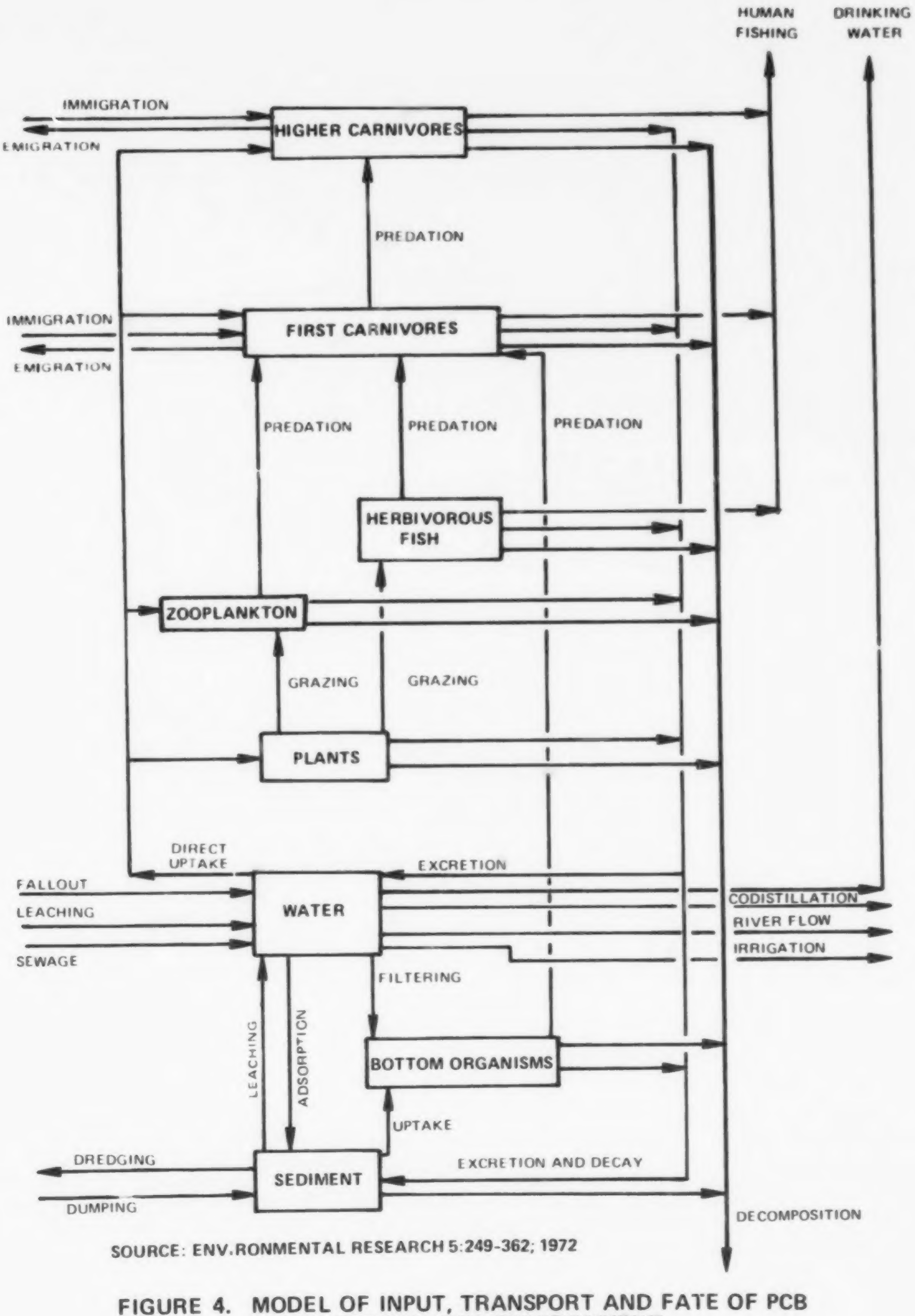
Models to describe sources, transport pathways and effects within global and regional environmental areas are needed to define monitoring schemes. An example of such a descriptive model is shown in Figure 4. Such models may have broad application for predicting movements and effects of many compounds that have similar properties and distribution in the environment. Perhaps more than any other form of pollution the problems associated with persistent synthetic compounds emphasize the importance of a conservative approach in pollution control.

## V. MULTIPLE EFFECTS OF POLLUTANTS

Although we often investigate the individual effects of pollutants on an organism, it is evident that effects in environmental systems are usually multiple and highly complex. Attempts to set maximum permissible levels of a pollutant must consider the full range of environmental interactions that may occur.

Temperature is an example of an environmental factor that has multiple effects on aquatic systems. Heated water discharge for electrical power plants is expected to increase in the next 30 years with substantial quantities of our surface waters used in the cooling systems. Individual heated discharges may have only minor effects on the environment but collectively several discharges in close proximity may create major environmental changes.

Individual aquatic species have specific thermal requirements that differ according to their life stages and influence their success in the environment.


HUMAN FISHING
DRINKING WATER
IMMIGRATION
EMIGRATION
HIGHER CARNIVORES
PREDATION
IMMIGRATION
FIRST CARNIVORES
EMIGRATION
PREDATION
PREDATION
PREDATION
HERBIVOROUS FISH
ZOOPLANKTON
GRAZING
GRAZING
PLANTS
DIRECT UPTAKE
EXCRETION
FALLOUT
LEACHING
SEWAGE
WATER
CODISTILLATION
RIVER FLOW
IRRIGATION
FILTERING
LEACHING
ADSORPTION
BOTTOM ORGANISMS
UPTAKE
DREDGING
SEDIMENT
EXCRETION AND DECAY
DUMPING
DECOMPOSITION

SOURCE: ENVIRONMENTAL RESEARCH 5:249-362; 1972

FIGURE 4. MODEL OF INPUT, TRANSPORT AND FATE OF PCB IN THE FRESHWATER ENVIRONMENT

Heated discharges can be managed to promote certain species if these requirements are known and protected. The maximum and minimum tolerance of desirable organisms may dictate the allowable differential between heated discharge and ambient water temperature. Thermal stress which is not lethal can have detrimental effects by increasing an organism's susceptibility to predation (Coutant, 1969). Each species also has some physiological optimum temperature for growth. However, as temperature increases food requirements also increase (Brett et al., 1969) and eventually a temperature is reached where growth ceases because of the high metabolic demand. Species that are forced to live near these marginal temperature extremes may be competitively excluded by species of greater thermal tolerance.

Fish in temperate climate zones require a seasonal fluctuation in temperature or winter chill period for successful development of the gametes and spawning. A period of low temperature is apparently required to allow storage of yolk material which cannot be effected successfully at higher temperatures because storage of food material is prevented at the higher metabolic rate (Merriman and Schedl, 1941). If fish are attracted to and remain in elevated water temperatures during the winter months they may suffer reproductive failure.

Other effects of temperature that may indirectly affect populations of aquatic organisms result from interaction with other pollutants. Macek (1969) reported insecticide toxicity to fish was increased at higher temperatures and Edsall and Yocum (1972) reported lake trout (*Salvelinus namaycush*) doubled their rate of methylmercury accumulation from water when the water temperature was increased from 5°C to 10°C. Pretreatment or exposure to a toxicant may reduce the tolerance to thermal stress in fish (Silbergeld, 1973).

An increased incidence of various fish diseases has been associated with elevated water temperatures in various locations (Mihursky, 1969; Brett, 1956). Ordal and Pacha (1967) found increased virulence and effect of several diseases on salmonids held at increased water temperature.

There is also some indication that heated water discharges may accelerate the process of eutrophication in lakes and create conditions favorable for nuisance algae (Patrick, 1969).

Temperature of course is only one of the many environmental variables that cause multiple or "synergistic" effects with other pollutants and different organisms. Nearly any pollutant can have multiple effects, many of which are not recognized until after harmful effects have occurred Wood, 1971 (cited by Bahr, 1972) reports the interaction of stream and lake enrichment by sewage effluent and the discharge of inorganic mercury as an example. The large bacterial populations created by sewage enrichment act as an "environmental catalyst" in the formation of the highly toxic methylmercury. Bacteria were probably always present in natural waters and so were small amounts of the element mercury. But increased discharge of elemental mercury in the presence of high bacterial populations allowed the accelerated production of methylmercury. Fish, birds and human populations have suffered from this unpredicted combination of environmental events.

## VI. POLLUTANT EFFECTS ON BIOLOGICAL COMMUNITIES

The laboratory testing of pollutant effects with important species has been recognized as a practical and necessary method for development of immediate pollution control standards. However, the limitations of the data obtained for a single species under highly artificial conditions is also recognized. An organism exists in nature as an integral part of the biological community, both exerting effects and being affected by interactions within the system. Improved methods for evaluating pollutant effects are needed to measure population or community responses to toxicants or other environmental perturbations. This is the next logical step for understanding the significance of effects observed at the species or individual organism level.

Warren and Davis (1971) have described methods of using laboratory stream ecosystems to measure dynamics of primary and secondary production of relatively simple communities of organisms. Even in these simplified systems the interactions between biotic and abiotic variables are complex, but much greater control can be exercised than in truly natural communities. The information obtained from these simulated communities or ecosystems offer a new level of sophistication for interpretation of environmental changes.

Maki and Johnson (1974) have used a series of model streams colonized by natural stream organisms to assess the effect of a toxicant TFM (3-tri-fluoromethyl-4-nitrophenol) on community structure and metabolism. TFM, which is used in streams to control populations of the parasitic sea lamprey (*Petromyzon marinus*), was introduced into streams under simulated field conditions. Their data show the toxicant caused immediate but reversible changes in the composition and abundance of sensitive invertebrate species and in the productivity of algal communities. Using radiolabeled TFM in separate sections of the stream they were able to measure the translocation and accumulation of residues in various organisms and the sediment. Although each of their measurements could have been conducted in isolated systems or with individual species, only by using a model community were they able to measure the rate of change and interaction of the different species.

Model aquatic ecosystems have also been employed to study the uptake and translocation of chemicals between several trophic levels. Isensee et al. (1973) studied the distribution of radiolabeled alkyl arsenicals in algae, snails, *Daphnia*, and fish in a simulated food chain. Their data provide information on the potential for bioaccumulation by the various species and transfer to higher trophic levels.

Each of the above examples is an attempt to improve on methods for assessing the effects of pollutants in the environment. Undoubtedly new, more precise methods will be developed. Although these methods will be more expensive initially the results should greatly improve our ability to predict and control environmental problems.

## VII. SUMMARY

Human society in its present form is highly dependent upon stable biological communities that form the basic structure of our life support systems. Our failure to recognize the integrated nature of ecological systems and man's dependence upon them has led to many of our present environmental problems.

Natural biological communities in a region consist of populations of organisms that have many common attributes. These similarities form a common denominator necessary for effective function and regulation of ecological systems. An immediate and practical need for pollution control is to recognize and protect the environmental requirements of individual populations of species within these regions. However, this must be considered a limited first step and much greater effort should be given to understanding the interrelationships between populations and communities.

Regional resource management plans offer an opportunity to view entire biological communities or ecosystems as functional units. Even at this level of organization we can err in management practices if we consider such regions as independent units. Ecological systems do not have distinct boundaries and change in one region will often effect changes in another.

Fish and wildlife populations are only one component of biological communities that we wish to maintain. However, in addition to their social and economic value, species of fish and birds often serve as important indicators of environmental quality. Their role and general productivity within a region are indicative of change that may be either to the benefit or detriment of human society.

## ACKNOWLEDGEMENTS

I extend my appreciation to the following people for their assistance in preparation of the manuscript: Dr. Harold Prince, Dr. Walter Conley, Mr. Alan Tipton, Ms. Judy Boger and Ms. Jackie Church.

## LITERATURE CITED

Alabaster, J. S., J. H. N. Garland, I. C. Hart and J. F. De L. G. Solbe. 1972. An approach to the problem of pollution and fisheries. Symp. Zool. Soc. Lond. No. 29: 87–114.

Adams, W. J. and H. H. Prince. 1972. Survival and reproduction of ring-necked pheasants consuming two mercurial fungicides. P. 306–317 *in* Environmental Mercury Contamination, R. Hartung and B. D. Dinman, *Editors*. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan, U.S.A. 349 pp.

Bahr, T. G., R. A. Cole and H. K. Stevens. 1972. Recycling and ecosystem response to water manipulation. Tech. Rept. No. 37: 124 pp. Institute Water Research, Mich. State Univ., East Lansing, Mich. 48824.

Borg, K., H. Wanntorp, K. Erne, and E. Hanko. 1969. Alkyl mercury poisoning in terrestrial Swedish wildlife. Viltrevy 6(4): 301–379.

Brett, J. R. 1956. Some principles in the thermal requirements of fishes. Quart. Rev. Biol. 31(2): 75–87.

Brett, J. R., J. E. Shelbourn, and C. T. Shoop. 1969. Growth rate and body composition of fingerling sockeye salmon, *Oncorhynchus nerka* in relation to temperature and ration size. J. Fish. Res. Bd. Canada 26(9): 2363–2394.

Brown, V. M. 1968. The calculation of the acute toxicity of mixtures of poisons to rainbow trout. Water Research 2: 723–733.

Cairns, J., Jr., R. E. Sparks, W. T. Waller. 1973. The relationship between continuous biological monitoring and water quality standards for chronic exposure P. 383–402 *in* Bioassay Techniques and Environmental Chemistry, G. E. Glass, Ed. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan, U.S.A.

Coutant, C. C. 1969. Temperature, reproduction and behavior. Chesapeake Science 10(3&4): 261–274.

Drummond, R A., W A. Spoor, and G. F. Olson. 1973. Some short term indicators of sublethal effects of copper on brook trout, *Salvelinus fontinalis*. J. Fish. Res. Bd. Canada 30: 698–701.

Duke, T. W., J. I. Lowe, and A. J. Wilson, Jr. (1970). A polychlorinated biphenyl (Aroclor 1254) in the water, sediment, and biota of Escambia Bay, Florida. Bull. Environ. Contam. Toxicol. 5: 171–180.

Eaton, J. G. 1973. Recent developments in the use of laboratory bioassays to determine "safe" levels of toxicants for fish. PP. 107–115, *in* Bioassay Techniques and Environmental Chemistry, G. E. Glass, *Ed*. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan, U.S.A.

Edsall, T. A. and T. G. Yocum. 1972. Review of recent technical information concerning the adverse effects of once-through cooling on Lake Michigan. Lake Michigan Enforcement Conference, September 19–21, 1972, Chicago, Illinois. 86 pp.

Fimreite, N. 1971. Effects of dietary methylmercury on ring-necked pheasants. Canadian Wildlife Service Occasional Paper No. 9. 39 pp.

Fimreite, N., R. W. Fyfe, and J. A. Keith. 1970. Mercury contamination of Canadian Prairie seed eaters and their avian predators. Canadian Field Naturalist 84(3): 269–276.

Ford, J. R. and H. H. Prince. 1972. Effects of methylmercury chloride on mallard reproduction. Presented at 1972 Annual Wildlife Disease Conference, June 13–15, University of Michigan, Ann Arbor, Mich.

Ford, J. R. and H. H. Prince. 1973. Distribution and excretion of methylmercury chloride in mallards. Presented at 35th Midwest Wildlife Conference, December 2–5, 1973, St. Louis, Missouri.

Friend, M. and D. O. Trainer. 1970. Polychlorinated biphenyl: Interaction with duck hepatitis virus. Science 170: 1314–1316.

Hansen, D. J., P. R. Parrish, J. I. Lowe, and P. D. Wilson. 1971. Chronic toxicity, uptake, and retention of Aroclor

1254 in two estuarine fishes. Bull. Environ. Contam. and Toxicol. 6(2): 113–119.

Herbert, D. W. M. 1962. The toxicity to rainbow trout of spent still liquors from the distillation of coal. Ann. Appl. Biol. 50:755–777.

Herbert, D. W. M. and D. S. Shurben. 1964. The toxicity to fish of mixtures of poisons. I. Salts of ammonia and zinc. Ann. Appl. Biol. 53: 33–41.

Herbert, D. W. M., D. H. M. Jordan and R. Lloyd. 1965. A study of some fishless rivers in the industrial midlands. J. Proc. Inst. Sew. Purif., pp. 569–582.

Isensee, A. R., P. C. Kearney, E. A. Woolson, G. E. Jones, and V. P. Williams. 1973. Distribution of alkyl arsenicals in model ecosystem. Environ. Sci. and Tech. 7(9): 841–845.

Jensen, A. L. 1971. The effect of increased mortality on the young in a population of brook trout, a theoretical analysis. Trans. Amer. Fish. Soc. 100(3): 456–459.

Jernelov, A. 1972. Mercury and food chains. P. 174–177 *in* Environmental Mercury Contamination, R. Hartung and B. D. Dinman, *Editors*. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan, U.S.A. 349 pp.

Johnson, H. E. and R. C. Ball. 1972. Organic pesticide pollution in an aquatic environment, pp. 1–10 *in* Fate of Organic Pesticides in the Aquatic Environment, Advances in Chemistry Series, No. 111.

Johnels, A. G. and T. Westermark. 1969. Mercury contamination of the environment in Sweden. P. 221–241 *in* Chemical Fallout, M. W. Miller and G. G. Berg, *Editors*. Charles C. Thomas, Springfield, Illinois. 531 pp.

Lloyd, R. 1961. The toxicity of mixtures of copper sulfate to rainbow trout (*Salmo gairdneri* Richardson). Ann. Appl. Biol. 49: 535–538.

Maki, A. W. and H. E. Johnson. 1974. Effects and fate of TFM (3-trifluoromethyl-4-nitrophenol) lampricide in model streams. Presented at 24th Annual Meeting, Midwest Benthological Society, Cincinnati, Ohio, March 27–29, Cincinnati, Ohio.

McKim, J. M. and D. A. Beneir. 1971. Effects of long-term exposure to copper on survival, growth and reproduction of brook trout (*Salvelinus fontinatis*). J. Fish. Res. Bd. Canada 28(5): 655–662.

McKim, J. M., G. M. Christensen, and E. P. Hunt. 1970. Changes in the blood of brook trout (*Salvelinus fontinalis*) after short term and long term exposure to copper. J. Fish. Res. Bd. Canada 27: 1883–1889.

Merriman, D. and H. P. Schedl. 1941. The effects of light and temperature on gametogenesis in the four-spined stickleback, *Apeltes quadracus* (Mitchell). J. Exp. Zool. 88(3): 413–449.

Mihursky, J. A., A. J. McErlean, and V. S. Kennedy. 1970. Thermal pollution in aquaculture and pathobiology in aquatic systems. J. Wildl. Dis. 6: 347–355.

Mount, D. I. and C. E. Stephan. 1967. A method for establishing acceptable toxicant limits for fish–mamathion and the butoxyethanol ester of 2, 4-D. Trans. Amer. Fish. Soc. 96: 185–193.

Nisbet, I. C. T. and A. F. Sarofim. 1972. Rates and routes of transport of PCB's in the environment. Environ. Health Perspectives 1:21–38.

Ordal, E. J. and R E. Pacha. 1967. The effects of temperature on disease in fish. *In* Water Temperatures: influence, effects and control. Proc. 12th Pac. Northwest Symp. Water Pollut. Res., Portland, Oregon. Fed. Wat. Pollut. Contr. Admin.

Panel on Hazardous Trace Substances Report 1972. PCB's–Environmental Impact. Environmental Research 5:249–362.

Patrick, R. 1969. Some effects of temperature on freshwater algae, pp. 161–185. *In* Biological Aspects of Pollution, P. A. Krenkel and F. L. Parker, *Editors*, Vanderbilt University Press. 407 pp.

Peakall, D. B. and J. L. Lincer. 1970 Polychlorinated biphenyls: Another long-life widespread chemical in the environment. BioScience 20: 958.

Ricklefs, R. E. 1973. Fecundity, mortality and avian demography, p. 366–435. *In* Breeding Biology of Birds, D. S. Farner, *Ed.* National Academy of Sciences, Washington, D.C. 515 pp.

Silbergeld, E. K. 1973. Dieldrin. Effects of chronic sublethal exposure on adaptation to thermal stress in freshwater fish. Environ. Sci. and Tech. 7(9): 846–849.

Spann, J. W., R. G. Heath, J. F. Kreitzer, and L. N. Locke. 1972. Ethylmercury p-tolune sulfonanilide: lethal and reproductive effects on pheasants. Science 175 (4019): 328–331.

Sprague, J. B. 1971. Measurements of pollutant toxicity to fish–III. Sublethal effects and safe concentrations. Water Research 5: 245–266.

Stalling, D. L. and F. L. Mayer, Jr. 1972. Toxicities of PCB's to fish and environmental residues. Environ. Health Perspectives 1: 159–164.

Stromborg, K. D. and H. H. Prince. 1974 Pesticide interactions: Alteration of reproductive performance in pheasants fed combinations of seed treatments. (In press).

Tucker, R. K. and D. G. Crabtree. 1970. Handbook of toxicity of pesticides to wildlife. U.S. Dept. of Interior, Bureau of Sport Fisheries and Wildlife Resource Publication No. 84. 131 pp.

U.S. Environmental Protection Agency. 1973. Effects of pesticides in water–a report to the States. USEPA, Washington, D.C. 20460. 145 pp.

Veith, G. D. 1972. Recent fluctuations of chlorobiphenyls (PCB's) in the Green Bay, Wisconsin, region. Environmental Health Perspectives 1: 51–54.

Warren, C. E. and G. E. Davis. 1971. Laboratory stream research: objectives, possibilities and constraints. Ann. Rev. Ecol. Syst. 1: 111–144.

# The Influence of Environmental Factors on Population Dynamics (Mathematical Model)

M. Ya. Antonovskiy *

1. We shall, in this report, discuss the following thesis. A population is a complex system containing a great many different heterogeneities that do, and do not, appear directly. At the same time, the observed, relatively stable, regime in which a population functions is the resultant of the effects of many factors, a number of which can cause instability.

The effect of the environment on a population can remove it from its stable regime and can bring into play internal mechanisms embedded in the population which can lead to the fundamental restructuring of the population, even when the magnitude of the external effect is small (more precisely, when there is a small change in the level of the effect of the environment on the population).

We shall, in this report, consider simple mathematical models of the different processes that can take place in "theoretical" populations that will, in our opinion, quite graphically illustrate the different aspects of the thesis we have set forth.

If we are to apply the results we obtain to real populations, it goes without saying that we must construct a model that will take into consideration the *specificity* of the interaction between the different factors in these populations. The appearance of this specificity can be expressed in the form of the corresponding quantitative dependencies constructed on the basis of a statistical analysis of the observed data. However, the difficulties that can arise when attempting to determine the precise form of the quantitative dependencies in biological systems are well known. Consequently, the approach to their study, which A. N. Kolmogorov [1] has defined as "obtaining quality results from quality prerequisites," is an extremely valuable one.

2. First of all, let us consider the phenomenon of the influence of the environment on the nature of the appearance of the internal mechanisms that depend on density, and which regulate the dynamics of population number and structure.

Study of the influence of density on population dynamics, experimentally, as well as using mathematical models, began a very long time ago.

The logistic equation

$$dN/dt = N(a - bN) \qquad (1)$$

was considered for the first time in connection with study of population dynamics by Fehrhulst in 1844, and led to the conclusion that there was a stationary level for the number to which the number in a population would tend monotonically. This fact is a classical example of negative feedback.

This conclusion retains its force, even when one must solve the more general equation

$$dN/dt = Nf(N) \qquad (2)$$

where

f(N) is a monotonically decreasing function, change in the sign of which has no effect on the conclusion reached as to the monotonic tending to some stationary level.

In fact, there are four types of number dynamics for populations functioning in a steady-state (in terms of time) environment that can occur (and they can be observed). These are:

(1) there is a steady stable state to which the system tends monotonically;

(2) there is a steady stable state to which the system tends in the form of attenuating oscillations;

(3) there is a stable, periodic, regime. At the same time, the magnitude of the period can be disassociated from the values for certain of the biological characteristics of the population (sexual maturation period, life span, and the like);

(4) the number undergoes irregular oscillations.

A discussion of this subject, as well as a great deal of experimental data, is contained in the basic monograph written by Andrewartha and Birch [2]. The book by Makfed'yen [3], while shorter, is full of information on this subject.

Population heterogeneity (and age differentiation in particular) is the most obvious candidate for the role of the cause of the effects listed.

However, at this point let us turn our attention to other possible causes. It may well be that the same population can experience all the types of dynamics listed, depending on the intensity of the effect of some external factor, such as a chemical or a physical agent, for example, that could have an effect on the fertility of females, or on the survival rate of specimens. Another example of an external factor is the degree of industrial exploitation of the population.

---

*Central Institute of Economic Mathematics, Moscow

This can be demonstrated on a model that does not take population heterogeneities into consideration.

Let us illustrate what we have said by a well known example. Let a population consist of specimens that have reached sexual maturity at year 1, and let the natural mortality and the fertility of the females be independent of age. Further, let the number of additions to the population per mature specimen be expressed by the monotonically decreasing function $\varphi(N,\mu)$. This function depends on the number of mature specimens, N, as well as on the parameter $\mu$, characterizing the influence of the environmental factor.

In this case the population dynamics equation has the form

$$N(t+1) = (1-\lambda)N(t) + N(t)\cdot\phi\left[N(t),\mu\right] \quad (3)$$

Here $\lambda$ is the mortality factor.

We shall not, here, dwell on the formulation of the general theorems that can be proven naturally with models of this type, but instead will consider the concrete, often encountered, example where

$$\phi(N,\mu) = \mu e^{-cN(t)} \quad (4)$$

where

c is some constant characterizing the ecological capacity of the environment.

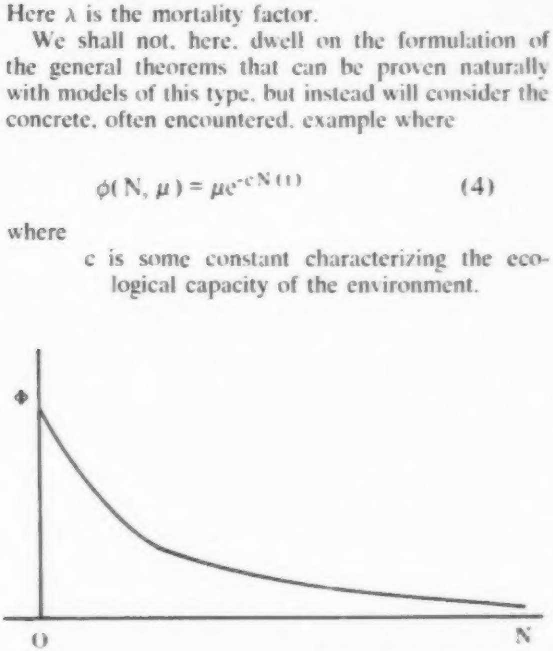


FIGURE 1. MODEL POPULATION FUNCTION

Eq. (4) makes it clear that the number of additions increases with increase in $\mu$.

Investigation of this model permits us to draw the following conclusions:

| | | | |
|---|---|---|---|
| 1 | $\mu \leqq \lambda$ | Extinction | |
| 2 | $\lambda < \mu \leqq \lambda e^{1/\lambda}$ | Monotonic growth to stable steady state | Case 1 |
| 3 | $\lambda + 1 < \mu \leqq \lambda e^{2/\lambda}$ | Tending to steady state with attenuating oscillations | Case 2 |
| 4 | $\mu > \lambda e^{2/\lambda}$ | Periodic or non-periodic regime | Cases 3 and 4 |

We should point out that when $\mu > \lambda e^{2/\lambda}$, the steady state will become unstable and the type of dynamics will depend on $\mu$, as well as on the original number in the population (as distinguished from cases 1 through 3). The type of dynamics will, for example, depend on the state in which the population was found when the parameter $\mu$ took a new value. There is a stable periodic regime for certain values of $\mu$ in which the population will be found for the "majority" of original conditions. Although the set of such original conditions for which the population will not be found in this regime is a small one, it has a measure of zero, the role of such conditions does manifest itself in the fact that in the case of other original conditions the reaching of the stable periodic regime is preceded by a more or less lengthy period (one that depends on the original conditions) of irregular oscillations "simulating" these unstable regimes. The length of the period will change in a most complex fashion with increase in $\mu$. There are certain $\mu$ values at which a phenomenon corresponding to case 4 will occur, and this we can call the ergodicity. The period of the stable cycle will become very large when the $\mu$ values are close to this, and periodicity will in fact disappear.

Let us consider in detail what happens when the parameter $\mu$ passes through the critical value $\lambda e^{2/\lambda}$. A trajectory with a period of $2(N_1, N_2, N_1, N_2, \ldots)$, which has the following properties, will appear in the neighborhood of the steady state $N = 1/c \ln \mu/\lambda$:

(a) it is locally stable; that is, for any original value N(O) sufficiently close to $N_1$, $N(2t) \rightarrow N_1$, $N(2t+1) \rightarrow N_2$, when $t \rightarrow \infty$;

(b) at the same time there are those original values N(O), the numbers for which N(t), do not tend to this stable trajectory; for example, $N(O) = 1/c \ln \mu/\lambda$;

(c) let us designate by $\triangle$ the set of those original values N(O) for which t exists such that

$$N(t) = 1/c \ln \mu\lambda.$$

Then $\triangle$ is an infinite denumerable set. The situation described in (b), above, will occur for any value N(O) of the closure $\overline{\triangle}$ of set $\triangle$. The trajectory will tend to the original trajectory of period 2 for any other original values. It is curious to note that the set $\overline{\triangle}$ is a "perfect Cantor set" of measure zero.

Similar effects are retained in more complicated models, in those closer to reality, in which the age structure of the population, change in fertility with age, and so on, are taken into consideration.

These models show that predetermined levels of effect of a hypothetical external factor are *critical* so far as the population is concerned; that is, the population will change its type of behavior when such value is reached.

There can be a situation in which the presence in the environment of an important ingredient, one however that is difficult to measure by using direct methods, can be established from the behavior of "indicator" populations that had been well-studied in laboratories, definite species of micro-organisms, or plants, for example.

We should point out that in recent years many of the published papers have considered population dynamics models based on type 4 relationships, or similar models with divisions into a small number of age classes. We can, by way of an example, point out reference [4], which deals with humpback salmon population, and in which a critical value was found at which the "natural" regime loses stability and in which computer data are cited for the case corresponding to the last row in our table.

3. Let us now go to the influence of the environment on a population that is the result of the effect on the genetic mechanism. The first phenomenon we shall analyze is the following. A genetically heterogeneous population in a steady (in some particular sense) state can be subjected to the effect of environmental factors that act on specimens of different genotypes in a differentiated manner. In this case the environmental factors act as a selection mechanism, and the rate at which this selection is made can prove to be extremely high. However, if, at the same time the influence of density on population dynamics still is taken into consideration, it can be shown that even the presence of a small advantage, obtained by a definite part of the population as a result of the effect of an environmental factor, can, with the passage of time, lead to a fundamental rearrangement of the genetic structure of the population (disappearance of certain genotypes, and the like).

Let us consider the following simple model in order to explain this thesis.

Let us assume the population to be genetically heterogeneous and consisting of two species (phenotypes). These phenotypes can be distinguished by a character that can be determined by one gene, found in two variants, a dominant A and a recessive a. Crossing of them takes place in accordance with the Hardy-Weinberg law, and each of the species experiences pressure from the number in the entire population, as described in paragraph 2, for example. We can assume, however, that the value of the parameter $\mu$ characterizing the effect of the environment is different for the two phenotypes, although both values are in the zone ensuring the existence of a stable steady state. Let us further suppose that $\lambda = 1$; that is, that each generation produces offspring once. This assumption is not critical for conclusions that can be reached using a model, but simplifies the calculations. If we designate the number of the total population at time t by $N(t)$, and the share of gene A in the genofund of the offspring produced by the generation living at time t as $x(t)$, then

$$x(t+1) = \frac{x(t)}{1 - \alpha\left[1 - x(t)\right]^2} \tag{5}$$

$$N(t+1) = \mu N(t) e^{-cN(t)} \left[1 - \alpha[1 - x(t)]^2\right] \tag{6}$$

where

$\mu_1 = \mu$ is the value of the environmental factor for the phenotype corresponding to genotypes AA and Aa;

$\mu_2 = \mu(1-\alpha)$ is the value of this factor for the $0 < < 1$ phenotype corresponding to genotype aa.

In both these numbers lie between 1 and 3, the net populations containing only genotype AA, or only genotype aa, have stable steady numbers, and the difference between them is small if $\alpha$ is small.

It may readily be seen that the mixed population too has a stable steady state:

$x = 1,\ 1/c \ln \mu,$

during which the recessive species disappears. But it is more interesting to estimate the rate at which this species is displaced.

Eq. (5) will provide this estimate quite readily. For example, if the original percentage of the species with genotype AA is equal to $p_0$, and is equal to $p_1$ after time n, we have

$$\frac{-\ln[1-\alpha(1-p_0)^2]}{-\ln[1 X \alpha(1 X p_0)^2)]} < n < \frac{-\ln[1-\alpha(1-p_1)^2]}{-\ln[1-\alpha(1-P_1)^2)]}$$

The second example of the influence of the environment on the population through the genetic mechanism that we shall analyze illustrating the theorem is that the genetic species can prove to be necessary in order to maintain the number of the population under variable (periodically changing, for example) environmental conditions.

It has long been established (references [6, 7, 8] for example) that there are clearly expressed differences between generations in populations of some small rodents, and other animal "ephemerals" that live for a few months and multiply several times in the summertime. The situation can be described, albeit very roughly, as follows.

Specimens with relatively low fertility predominate in a generation that overwinters and has offspring at the beginning of the season. However, a generation, for example, that is born in the first half of the summer ("descendants" of the previous generation) is predominated by specimens with a very short sexual maturation period, is very fertile, and has a short life span. Since the population is substantially reduced during the wintertime, the population is restored to its former figure. It is considered established, at least for some populations, that the difference between the species we have described ("living through the winter" and "summer") is in fact genetic in nature.

We have been considering a very simple mathematical model that suggests one possible approach to an explanation for the need for periodic rearrangement of the genetic structure of a population in order for that population to survive.

Just as was the case with the preceding models, we must bear in mind the fact that our purpose simply is to present a graphic, mathematically correct demonstration of a definite phenomenon by the example of a simple "theoretical population." We should note that a number of the simplified assumptions can be discarded without creating any difficulties. There are others that are more significant in content, however, and which we shall discuss in what follows, that must be taken into consideration, even in the more complex models, those in which the effect of the influence of polymorphism on support of the number in the population also can arise.

Let the population under consideration consist (as in the preceding example) of two phenotypes determined by one gene with dominant (A) and recessive (a) alleles.

Let us also assume that two reproduction cycles take place during a season, and let us, for purposes of simplicity, assume that each generation reproduces but once. The generation born during the second reproduction cycle will not reproduce in this same year, and the part of this generation that overwinters is the first of the reproducing group the following year.

Let $\alpha_1$ be the mean number of descendants among the specimens with genotype AA and Aa, and $\alpha_2 = \beta\alpha_1 (\beta > 1)$ be the mean number of descendants of the specimens with genotype aa.

Let us make the further assumption that there will be no deaths among those specimens in the population that do not reproduce during the summertime, and that during the fall and winter $\lambda_1$ percent of the number of specimens with genotypes AA and Aa, and $\lambda_2$ percent of the number of specimens with genotype aa, will die; $0 < \lambda_1 < \lambda_2 \leqslant 1$.

Accordingly, the specimens with genotypes AA and Aa (let us call them the "winter" specimens) have a relatively low fertifility rate, but tolerate the fall and winter better. Those specimens with the "summer" genotype aa are more fertile, but at the same time are more sensitive to environmental changes.

We shall here, as in the preceding model, take it that panmixia takes place in the population.

Now let X, Y, and Z be the numbers of specimens with genotypes AA, Aa, and aa, respectively. $N = X + Y + Z$ is the total number of the population, and

$$x = X/N, \; y = Y/N, \; z = Z/N$$

are the percentages of the corresponding genotypes.

The vector (X, Y, Z) determines the structure of the population at each moment in time. If the entire generation lives until the next propagation period the structure of the succeeding generation can be determined by the operator T

$$T(X, Y, Z) = (X', Y', Z')$$

where

$$X' = \frac{\alpha_1 (X + 1/2Y)^2}{N};$$

$$Z' = \frac{(\alpha_1/2 \; Y + \alpha_2 Z)(Z + 1/2Y)}{N};$$

$$Y' = \frac{\alpha_1 (X + 1/2Y)(Z + 1/2Y)}{N}$$

$$\frac{+ (\alpha_1/2 \; Y + \alpha_2 Z)(X + 1/2 \; Y)}{N}.$$

Let us further assume that

$$S(X, Y, Z) = [(1 - \lambda_1)X, (1 - \lambda_1)Y, (1 - \lambda_2)Z].$$

Then the population with the structure (X, Y, Z) at the beginning of the summer will have the structure

$$(ST^2)\ (X, Y, Z) = (X_1, Y_1, Z_1)$$

at the beginning of the next summer.

It can safely be pointed out that the corresponding transformation of the variables (x, y, z) (for which $x + y + z = 1$) is completely determined by their values (that is, it does not depend on N). The analytical form of the representation of the triangle used in this case is much more complex than in the conventional iterations of quadratic representations in mathematical genetics (see reference [8]). Formally, the reason for this fact reduces to this. One is forced to make a certain normalization, as distinct from a purely genetic situation, in which the sum of the coordinates can be preserved automatically (even if panmixia is lacking) in order to retain the condition that $x + y + z = 1$ because of the differences in fertility and mortality among the different genotypes.

Obviously then, a pure population consisting only of specimens with genotype AA will survive if

$$a_1{}^2(1 - \lambda_1) \geqslant 1 . \qquad (7)$$

Analogously, a pure "summer" population will survive when

$$a_2{}^2(1 - \lambda_2) \geqslant 1. \qquad (8)$$

We shall demonstrate the fact that a mixed population, in which "winter" and "summer" specimens are present, will survive in the case of certain values of the parameters $\alpha_1$, $\alpha_2$, $\lambda_1$, $\lambda_2$, that satisfy neither of the inequalities at (7) and (8).

We should point out that in the model under consideration the dependence of fertility on density has been omitted for reasons of simplicity. However, if this dependence is introduced, as was done in the preceding model, similar effects will occur here as well.

We have made yet another simplification, one that allows us to avoid making cumbersome calculations. To that end, let us assume that $\lambda_2 = 1$ (the "summer" specimens all die during the winter). This of course is not compatible with the inequality at (8), so in this case it is convenient to use the variables N and $\eta = y/2$.

Let us put $\alpha_1 = \alpha$, $\lambda_1 = \lambda$, $\alpha_2 = \beta\alpha$. Direct calculations will lead to the following relationship, one determining population dynamics during the year

$$(N, \eta \rightarrow (N_1, \eta_1)$$

where

$$N_1\ (N, \kappa) = (1 - \lambda) N a^2 (1 - \kappa)\ [(1 + \kappa - \kappa^2) + \beta\kappa^2)] \qquad (9)$$

$$\kappa_1\ (\kappa) = \frac{\kappa\,[2 + (\beta - 1)\kappa]}{2 + 2\kappa + (2\beta - 2)\kappa^2} \qquad (10)$$

The value $\eta = 0$ is a fixed point on the representation of the segment [0, ½] itself, found by using Eq. (10). This point is stable when $\beta < 3$, with the result that in a mixed population of heterozygous specimens (and, consequently, the percentage of summer specimens that appear during the season) it vanishes. In this case the conditions for the survival of the mixed population are given by the inequality at (7), just as was the case for the pure population.

However, when $\beta > 3$, the point $\eta = 0$ will become unstable, and another stable, fixed point $\eta = (\beta - 3)/2(\beta - 1)$, will appear (see Figure 2).

What this means is that each spring the population will contain a definite percentage of heterozygous specimens, so that summer specimens will appear in the course of the reproduction season, and will reproduce. These specimens represent a not inconsiderable part of the population. The condition for the survival of a mixed population, readily obtained from the inequity $N_1(N, \eta) \geq N$, has the following form when $\beta > 3$

$$a^2 \left[1 + \frac{(\beta - 3)^2}{8(\beta - 1)}\right](1 - \lambda) \geqslant 1 \qquad (11)$$

Thus, if for given $a$ and $\beta$

$$a^2 \leqslant \frac{1}{1 - \lambda} \leqslant a^2 \left[1 + \frac{(\beta - 3)^2}{8(\beta - 1)}\right],$$

the pure "winter" population will die out, but the mixed population will not! (Let us recall that the pure "summer" population too will die out, because $\lambda_2 = 1$.)

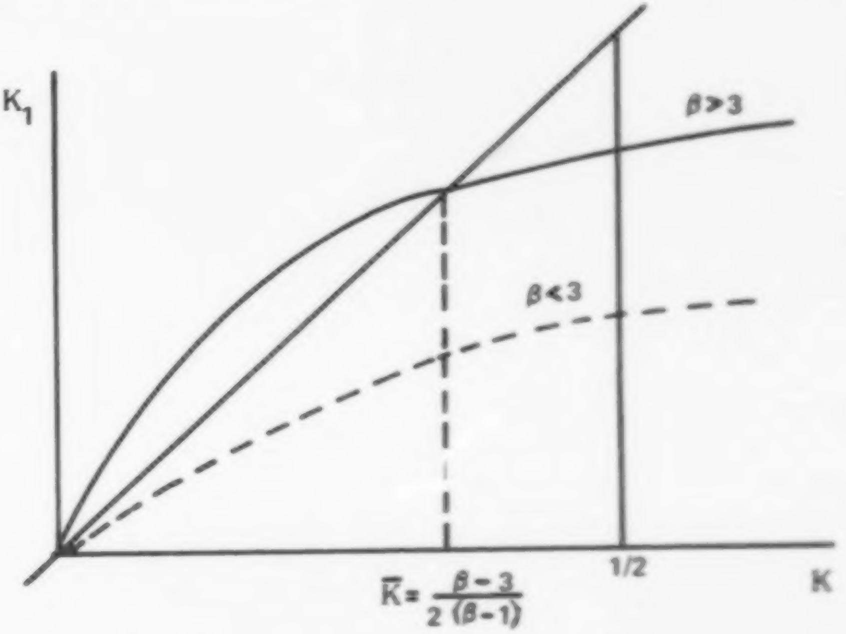


FIGURE 2. POPULATION DYNAMICS

The coefficient $1+(\beta-3)^2/8(\beta-1)$ measures the advantage in viability imparted by polymorphism.

Let us use a numerical example that will show quite graphically the population dynamics in this case to illustrate our conclusions.

Let $\alpha=2$, $\beta=5$ (that is, $\alpha_2=10$), $\lambda=7/9$. In this case $\mu_0=½$. Let us set $X=Y=100$.

| Times | t | t+1 | t+2 | t+3 |
|---|---|---|---|---|
| X | 100 | 225 | 450 | 100 |
| Y | 100 | 150 | 450 | 100 |
| Z | 0 | 25 | 100 | 0 |
| N | 200 | 400 | 1000 | 200 |

1 year

For purposes of comparison with a pure "winter" population

|  | t | t+1 | t+2 | t+3 |
|---|---|---|---|---|
| X=N | 200 | 400 | 800 | 178 |

We have selected limit $\lambda$ values for which the inequality at (11) will be transformed into an equality, despite the fact that from the standpoint of our model this is an exceptional case, in that a small change in $\lambda$ will lead to dying out, or to exponential growth of N and to increasing oscillations in the magnitudes X, Y, Z. However, this behavior will be typical of a model in which fertility is a function of density.

We shall conclude our discussion of this model with a few remarks. Along with the simplifications already mentioned, we have assumed that one gene has a simultaneous influence on fertility and survivability under severe environmental conditions. This assumption would not appear to be very natural at first glance. However, it can be supposed that the "winter" and "summer" species can be distinguished by metabolic intensity, the consequence of which is a longer period of sexual maturation and a longer life span among the "winter" specimens.

Here consideration must be given to the fact that the greater fertility of the "summer" specimens is the result of their success in reproducing a greater number of times.

A scheme such as this can be realized in a model if, needless to say, the number of reproduction cycles per season is increased (for the "summer" specimens at the very least). Another extremely interesting direction is associated with study of the oscillations that take place in the genetic structure in a population of long-lived organisms (see reference [9]).

4. The final subject we shall touch upon in this report is the effect man has on a population as it is exploited; that is, by the withdrawal of some of the specimens from the population. The practical importance of this subject is obvious, and its study covers a considerable part of mathematical ecology (see the well-known book by K. Watt [10], for example, in which the author concentrates on the question of the effect exploitation has on a population).

We should like to turn our attention to the connection that exists between the subject of optimum exploitation of a population and the principal theorems in mathematical economics.

Let us consider the following model. Let a population consist of N+1 mature groups. Let us designate the number in the $i^{th}$ group by $n_i$, $i=0, 1, \ldots, N$. $F_i$ is the mean number of offspring per unit time for the specimens in the $i^{th}$ group. $p_i$ is the probability that specimens aged i will live to age i+1 under natural conditions (no exploitation).

Let us, in addition, introduce the index $W_i$, the exploitational worth of the specimen.

Let us give our external effect on the population (or control) by the vector $(u_0, u_1, \ldots, u_N)$, where $u_i$ is the percentage of the $i^{th}$ group meaningfully withdrawn from the population (for the time segment under consideration).

We must, in order to establish population and harvest dynamics, stipulate the moment in time when the harvest must be removed (this need not be done if a continuous time model is considered).

If the harvest is removed at the beginning of the period we shall call the corresponding model *model A*. It has the form

$$\begin{cases} n'_0 = (1-u_0)\sum_{i=0}^{N} F_i n_i \\ n'_i = (1-u_i)p_{i-1}n_{i-1},\ i \geq 1 \end{cases} \tag{12}$$

The harvest is equal to

$$W = \sum_{i=0}^{N} \frac{u_i}{1-u_i} \cdot n_i \cdot W_i$$

587

($W_i$ can obviously be equated to the beginning of the period as well.)

If we remove the harvest at the end of the period (and after taking the effect of natural mortality into consideration), we shall call the corresponding model *model B*.

$$\begin{cases} n'_0 = \sum_{i=0}^{N} F_i n_i \\ n'_i = (1-u_{i-1})p_{i-1}n_{i-1},\ i \geq 1 \end{cases} \tag{13}$$

The harvest is equal to

$$W = \sum_{i=1}^{N} u_{i-1}p_{i-1}n_{i-1}W_{i-1} = \sum_{i=0}^{N-1} u_i p_i n_i W_i \quad (14)$$

($W_i$ in this case can be equated to the end of the period.)

We shall, for purposes of definiteness, henceforth consider the type A model.

The problem, in the most general of terms, is that of constructing that sequence $u^{(1)}, u^{(2)}, \ldots,$ of equations that will maximize the harvest.

We have to know the factors upon which the dynamics of the number in a population depends in order to state the problem more precisely (density, nutrition, and the like).

It can be taken, for example, that the parameters $F_i$, $p_i$, $W_i$ depend on n in such a way that an increase in n will reduce their values.

Another approach is that there is some set of reserves and that a limitation of the following type can be satisfied

$$\sum \alpha_{ij} n_i \leqslant K_j \quad (j=1,2\ldots,m) \quad (15)$$

where

$\alpha_{ij}$ is the consumption of reserve j by one specimen in the $i^{th}$ group;

$F_i$, $p_i$, $W_i$ are constants that still satisfy the limitation at (15).

Then we must clearly distinguish what it is that we take as the criteria. If expenditures required to remove the harvest are ignored, we can, for example, maximum the summed harvest over a long period of time.

We should point out that we must, naturally, limit those controls that depend solely on the present state of the system.

Employing the considerations that can be used for the proofs of the so-called theorems of main lines in consumption (see reference [11], Chapter IV), we can prove the following. *There exists that temporally independent steady-state equation* $u = u_0, u_1, \ldots, u_{N-1}$, *and the vector n, invariant with respect to this equation, that is, satisfying the condition set forth at (16)*

$$\begin{cases} n_0 = \sum_{i=0}^{N} F_i n_i \\ \\ \\ n_{i+1} = (1 - u_i) p_i n_i, \; i = 0, \ldots, N-1 \end{cases} \qquad (16)$$

*and any optimum trajectory of length T always will reduce to a small neighborhood of state n, except for the end segments at the beginning and at the end, where the controls used at this time are close to control u.* We will not, here, discuss in detail the occurrence of the analogy between the economic and the ecological magnitudes. We shall simply note that the limitations of (15) play the same role that the exogenous reserves play in an economic system; that is, those reserves the system itself does not condition (manpower, for example).

The considerations cited, which are new only in part, have been advanced as a result of discussions between the participants in the group engaged in the mathematical modeling of ecological systems in the Laboratory of Functional Analysis and Models of Economic Dynamics at the TsEMI [expansion unknown] of the Academy of Sciences of the USSR. The group includes M. Ya. Antonovskiy, V. I. Danilov, V. A. Iskovskikh, and A. B. Katok.

Academician S.S. Shvarts raised for our group questions concerned with the interaction of ecological and genetic factors. The phenomena he pointed out are at the basis of the mathematical models considered in paragraph 3.

Discussions of the questions related to the influence of density on population dynamics with our coworkers at the Institute of the Ecology of Plants and Animals of the Academy of Sciences of the USSR, and with V. S. Smirnov, and V. G. Ishchenko in particular, were extremely useful.

## BIBLIOGRAPHY

1. A. N. Kolmogorov. "Qualitative Investigation of Mathematical Models of Population Dynamics." *Problemy kibernetiki*, No. 25, 1972, pp. 100-106.
2. H. G. Andrewartha, L. C. Birch. *The Distribution and Abundance of Animals*. The University of Chicago Press, Chicago, 1954.
3. E. Maklefyen. *Ekologiya zhivotnykh* [The Ecology of Animals]. "Mir" Publishing House, Moscow, 1965.
4. V. L. Andreyev, Yu. A. Nagorskiy, A. P. Shapiro. "Modeling a Fish Population With a Two-Year Life Cycle and a Single Spawning." *Problemy kibernetiki*, No. 25, 1972, pp. 167-176.
5. H. D. Ford, E. B. Ford. "Fluctuation in Numbers and Its Influence on Variation in *Melitaea Aurinia*." *Ent. Trans. Royal Ent. Soc.*, London, Vol. 78, 1930.
6. S. M. Gershenzon. "Seasonal Changes in the Frequency of the Occurrence of Black Hamsters." *DAN SSSR*, Vol. 43, No. 8, 1945.
7. S. S. Shvarts. *Osnovy evolyutsionnoy ekologii zhivotnykh* [Fundamentals of the Evolutionary Ecology of Animals]. Sverdlovsk, 1969.
8. F. Frank. "The Causality of Microtine Cycles in Germany." *J. Wildlife Man.*, 21, 115-121 (1957).
9. S. S. Shvarts, V. G. Ishchenko. "Dynamics of the Genetic Composition of a Population of *Rana arvalis*." *Byulleten' Moskovskogo obshchestva ispytateley prirody, Otd. biologii*, Vol. 43, No. 3, 1968, pp. 127-133.
10. K. Watt. *Ekologiya i upravleniye prirodnymi resursami* [Ecology and Control of Natural Resources]. "Mir" Publishing House, Moscow, 1971.
11. Kh. Nikaydo. *Vypuklyye struktury i matematicheskaya ekonomika* [Convex Structures and Mathematical Economics]. "Mir" Publishing House, Moscow, 1972.

# Global Climate and Human Activity

M. I. Budyko,* I. L. Karol *

## 1. Models of Changes in Global Climate

Previous investigations (Budyko, 1969, 1971, 1972; SMIC 1971, and others) have shown that modern economic activity already has had a marked influence on global climatic conditions, and this influence is increasing rapidly.

This fact has given rise to the important task of estimating possible climatic changes in the future for different conditions of further economic development.

The only way to accomplish this task is to use numerical models of climate theory to determine changes in meteorological conditions.

Models of a climate investigated for purposes of studying its possible future changes should satisfy certain requirements, but these latter are less important for models used in investigations of the present day climate.

The model ought not include empirical data on the distribution of individual climatic elements, especially those which change significantly with changes in climate.

The model should take into consideration all types of heat flows with a marked effect on the temperature field, and the law of the conservation of energy should be satisfied.

Climate is formed by:

1. the distribution of radiation in the atmosphere;
2. the transfer of heat and moisture by the movements of the atmosphere and ocean;
3. the hydrology of the underlying surface (it has a greater influence on the local climate of the lower troposphere, and is not considered here).

The heat balance equation

$$Q(1-A)-I=F, \qquad (1)$$

links the flow of heat from the sun's shortwave radiation, Q (A is the albedo of the earth-atmosphere system), the infrared (IR) radiation of heat, I, for this system, and the heat transfer, F, by the movements of the atmosphere and ocean, as basic factors. These factors depend significantly on the following parameters (Budyko, 1971): A, the albedo of the earth-atmosphere system, increases sharply over (a) a snow and ice surface, (b) clouds (especially dense clouds), (c) without clouds somewhat more over land bare of vegetation than over the sea.

The following also have a great effect on the distribution of radiation in the atmosphere: (a) clouds and water vapor absorbing and scattering short- and longwave radiation; (b) dust, doing the same to the sun's shortwave radiation; (c) carbon dioxide, $CO_2$, which absorbs IR radiation; (d) ozone, $O_3$, which heavily absorbs the sun's ultraviolet (UV) radiation.

Global movements of the atmosphere are determined by the planetary distributions of its temperature and humidity and, in turn, form them. Their intensity depends primarily on the difference in temperatures at the equator and at the poles, and between the continents and oceans. The relationship between the nature and intensity of ocean currents and external conditions is more complex. It is only partially linked with atmospheric circulation.

These climatic factors are linked by the following positive and negative feedback rings.

1. *Polar ice*, almost doubling A, reduces the air temperature, T, and thus promotes its own distribution. This is a positive feedback ring, which exerts a great influence on the distribution of T (Budyko, 1971).

2. *Clouds*, generally speaking, reduce the radiation passing to the earth, and reduce T, and evaporation, E, which leads to a reduction in cloud cover. This is a negative feedback ring. But they do inhibit IR radiation and block and radiation reduction in T. This is a positive feedback ring.

3. *Increase in absolute humidity* with increase in T leads to the absorption of longwave radiation by the atmosphere and to its heating. This is a positive feedback ring.

4. *Increase in T* increases the intensity of the atmosphere's IR radiation. This is a negative feedback ring.

5. *Increase in the $CO_2$* in the atmosphere reduces the intensity of its absorption by the ocean, given a rise in temperature and in the acidity of the waters of the latter. This leads to an even greater accumulation of $CO_2$ in the atmosphere, and is a positive feedback ring (Walker, 1973).

---

*Main Geophysical Observatory, Leningrad

6. *The intensity of heat and moisture transfer* decreases with warming of the polar atmosphere, and so does the transfer of heat from the equator to the poles. This is a negative feedback ring.

Negative circulation rings increase climatic stability; positive ones decrease it.

The parameters of virtually all of these feedback rings have not yet been adequately studied, and it is possible that other feedback rings exist, but if they do they are not significant.

Man can affect climate in the following ways.

(1) A change in the components of the heat balance of the underlying surface, of a city, a reservoir, irrigation of fields and of agricultural land, is the most noticeable, and comparatively well-studied change in climate, but is only local, and in the lower troposphere.

(2) The emission of heat into the atmosphere even now is changing local climate in cities ("heat islands" in the center of a city). When the current rate of growth in the world's energy requirements (about 6% a year) is extrapolated, the magnitude of this emission in the 21st century will become comparable to the magnitudes of the components of the heat balance for the underlying surface, particularly in urban and industrial agglomerations (Budyko, 1972).

(3) The increase in the $CO_2$ in the atmosphere attributable to the combustion of fossil fuel in recent decades is 0.2% per year (Machta, 1973). According to the estimates contained in this reference, the mean concentration of $CO_2$ in the atmosphere in 1860, was approximately 10% lower, and in the year 2000, will be 15–20% higher, than its current value.

(4) The increase in the content of atmospheric aerosols. Stratospheric aerosols on the whole reduce solar radiation incoming to the lower atmosphere and the underlying surface. Heating of the stratosphere by radiation absorbed by these aerosols has almost no influence on the temperature of the troposphere, and thus as a practical matter reduces the meteorological solar constant (Kondrat'yev, et al., 1973).

The effect of tropospheric aerosols is more complex. They scatter and absorb radiation, and this respectively increases and decreases the albedo (A) of the "earth-atmosphere" system. This change in A depends on the effective albedo of the aerosols, $A_a$, and on the albedo of the underlying surface, $A_s$. If $A_s$ is small, and if $A_a > A_s$ (over a cloudless ocean), then an increase in the aerosol content, C, increases A and reduces the temperature, T, of the lower atmosphere. If $A_s > A_a$ (over snow and clouds), then an increase in C reduces A and increases T. $A_a$ depends on the aerosol's optical properties and can change greatly in different regions, and in different seasons, but as a mean, $A_a$ is close to the mean for $A_s$ in the absence of clouds, snow, and ice (Kondrat'yev, et al., 1973). Aerosols under a cloud layer have little influence on the radiation flux near the earth.

Aerosols of anthropogenic origin now comprise 10–20% of the total mass of atmospheric aerosols. There are predictions that suggest a doubling of this percentage by the year 2000, if special measures are not taken to reduce the emission of aerosols into the atmosphere by man (SMIC, 1971).

(5) Mass flights of aircraft at a level of about 20 km can result in the accumulation in the stratosphere of the combustion products exhausted from their engines, in a decrease in the ozone content by interaction with the nitrogen oxides they emit, and in an increase in stratospheric aerosols (Grobecker, 1973). The possible effect of these aerosols on climate was indicated above, and a reduction in the ozone layer will lead to an increase in the biological danger from the sun's UV radiation. A change in the ozone content can change the climate in the stratosphere (especially its circulation), because the absorption of solar radiation by ozone to a considerable degree determines the heating and temperature of the stratosphere.

Various assumptions can be made with respect to each of these factors, on how each will change in the future as further economic development takes place. The task of climatological investigations is to estimate climatic conditions for possible variants of change in these factors. The question of the reliability of estimates of these changes is worthy of serious attention because of the great practical importance of global climatic changes.

Two different approaches can be used in considering this question. The first involves the verification of the numerical models used for calculating the climate of the future using empirical data on climatic changes in the past. This approach was used in previous work by one of the authors (Budyko, 1971, 1972, and others).

The second approach is based on a comparison of the results of calculations of climatic conditions of the future, using different, independently developed models of climatic theory. The feasibility of using this approach appeared recently as a result of the making of a series of investigations of anthropogenic climatic changes.

Let us consider the results of calculations of the influence of increasing the production of energy on climate, made using various models of climatic theory.

Washington (1971) used a three-dimensional model of climatic theory to calculate the change in their temperature field in the troposphere as the result of an additional heat inflow of 50 cal/cm²/day, equally distributed over the surface of the continents.

The mean air temperature in this case rose from 1° to 2° C in the tropics, and from 8° to 10° C in the higher latitudes of the northern hemisphere.

A calculation using a semiempirical model of climatic theory (Budyko, 1969) yields approximately the same increase in the mean global temperature of 3° C for these conditions (note that in neither case was the dependence of the albedo of the earth-atmosphere system on temperature changes taken into consideration).

Estimates of change in the carbon dioxide concentration with thermal conditions in the atmosphere were made in our paper (Budyko, 1973) and in that by Manabe (SMIC, 1971; Smagorinsky, 1974).

Let us note that the feedback between changes in thermal conditions in the atmosphere and the position of the snow and ice cap were taken into consideration in all these studies.

A comparison of the results of the calculation of the effect of doubling the concentration of carbon dioxide on the mean latitudinal air temperature, made using our semiempirical climatic theory, and the general climatic theory used by Manabe, shows good agreement between the estimates of the influence of change in carbon dioxide concentration on climate obtained using the different models. Hence, it follows that in such case the mean air temperature increases by 2°–3° C in the low latitudes, and by up to 10° C in the high latitudes.

The question of the influence of an aerosol on thermal conditions in the atmosphere has been discussed in numerous studies. It is determined by the radiation flux in the atmosphere as a function of the aerosol concentration, and by the air temperature as a function of changes in shortwave radiation flux.

Leaving the first part of this link aside for the moment, let us consider the results of the calculations of air temperature as a function of the inflow of solar radiation. These calculations have been made repeatedly for different models of climatic theory without taking into consideration the feedback between thermal conditions in the atmosphere and polar ice.

This feedback was taken into consideration for the first time within the framework of a semiempirical model of climatic theory (Budyko, 1969). This study established that the mean air temperature at the earth's surface changes sharply with small fluctuations in the solar constant, and that a reduction of approximately 2% in the solar constant will result in complete glaciation of the earth. A similar conclusion was obtained as a result of the development of a series of different semiempirical models of thermal conditions in the atmosphere, including consideration of the feedback indicated above (Sellers, 1969; Schneider and Gal-Chen, 1973).

The conclusion drawn with respect to the high degree of the sensitivity position of the polar ice to small changes in the inflow of solar energy resulted in obtaining a quantitative explanation of the process of the development of quaternary glaciation (Budyko and Vasishcheva, 1971; Berger, 1973).

The question of air temperature as a function of the solar constant was studied in the last papers written by Manabe (Smagorinsky, 1974) who used general climatic theory and included consideration of the feedback between thermal conditions and the snow and ice cap. Here, it was found that a 2% increase in the solar constant resulted in a 1.2° C mean change in the air temperature for the hemisphere, according to the Manabe and Vezerold local model of thermal conditions when changes in absolute air humidity were disregarded. The change was 2.4° when absolute air humidity as a function of temperature was included and the same model was used, but was 3.1° C when Manabe's model of general climatic theory was used. The latter magnitude differs comparatively little from the result of the analogous calculation made using the above-mentioned semiempirical model of the thermal conditions in the atmosphere, from which it follows that the indicated increase in temperature is 4.5° C.

In estimating the magnitude of the relative reduction in the solar constant sufficient to completely glaciate the planet, Manabe, as well as the authors of the semiempirical studies, found it comparatively small, and approximately equal to 4%.

Thus, a great many studies based on the use of totally different models of climatic theory produce what are comparatively close estimates of the influence of a change in climate-forming factors on climate. This permits us to use these models with a known degree of confidence in making calculations of climate changes that could result from the development of economic activity.

## 2. Changes in Climate by the End of the 20th Century

Let us examine the climatic changes that could take place by the end of our century as a result of anthropogenic factors. Calculations suggest that increase in energy production (by approximately 6% a year) will have no significant influence on climate before the end of the 20th century, although this influence could be very substantial over the next century (Budyko, 1972).

Calculations of the influence of aerosols on thermal conditions in the atmosphere are difficult to make because of the lack of data on the mean values of the parameters of the models used, particularly as the data pertain to the coefficients of radiation absorption by aerosol particles. Here it is that analysis of empirical data is of considerable importance in estimating the influence of an aerosol on radiation and thermal conditions.

The data from observations made by a group of actinometric stations in Europe, Asia, and America were processed by Z. I. Pivovarova and the results were used to chart the secular course of direct solar radiation in a cloudless sky over the past decade for purposes of estimating the influence of an anthropogenic aerosol on climate (Budyko and Vinnikov, 1973). The conclusion reached as a result of this investigation was that anthropogenic aerosols have in recent years reduced direct radiation in a cloudless sky by approximately 6%, and the summed radiation for average conditions by approximately 0.5%.

The use of the Yamamoto and Tanaka (1972) model, in which the influence of aerosol on backscattering and radiation absorption is taken into consideration, resulted in the finding that this reduction in radiation corresponds to a reduction in the mean air temperature at the earth's surface of approximately 0.5° C.

Although, as noted in this study, this estimate is very approximate in nature, it can be assumed that further rise in the aerosol concentration can lead to a significant change in climate. However, it is quite difficult to predict the future quantity of atmospheric aerosols.

We can assume that failure to combat atmospheric pollution will result in the ratio of increase in the anthropogenic aerosol mass to its current value being equal to the analogous value for the anthropogenic carbon dioxide mass in the atmosphere. Given this assumption, the quantity of anthropogenic aerosol can increase by a factor of 2.5 to 3 by the year 2000. But because many countries are taking steps to reduce atmospheric pollution, the increase in the quantity of anthropogenic aerosols will come to a halt, and their quantity by the end of the century will remain at the current level, or may even decrease. Given these conditions, increase in the carbon dioxide concentration, leading to a warming trend, will exert the major influence on change in climate.

A variety of models of climatic theory can be used to estimate this warming trend. What follows from the above-cited Manabe and Vezerold local model of climatic theory is that a 20% increase in carbon dioxide concentration will increase the mean global air temperature 0.5°. Manabe's calculations, based on the use of a three-dimensional model of general climatic theory, showed a 0.7° C increase in the mean global temperature for this case, and approximately twice that for the high latitudes.

A similar calculation, using our semiempirical model of thermal conditions in the atmosphere and the relationship between air temperature and carbon dioxide concentration according to Rakipova-Vishnyakova, and given the assumption of the quasi-stationary state of polar ice (see Budyko, 1972), showed a 0.6° C increase in temperature, and, given the assumption of stationary ice conditions, 0.9° C. As we see, there is little difference between these estimates.

Calculations using a semiempirical model of thermal conditions in the atmosphere show that the area of sea ice in the northern hemisphere can be reduced by 10 to 20 percent by an increase of 0.5° to 1° C in the mean global temperature. This will be accompanied by a several degree increase in temperature during the winter in the high latitudes, something that can greatly influence Arctic navigation conditions, as well as many branches of the national economy in the northern regions.

Fluctuations in the amount of precipitation that falls, associated with change in thermal conditions, can be of even greater practical importance. The overall warming trend can be accompanied by a reduction in the mean meridional temperature gradient, because the air temperature in the high latitudes rises much more than it does in the low latitudes with increase in the carbon dioxide concentration. As is known, changes in this gradient have very definite influence on the intensity of zonal circulation.

A series of empirical studies (Budyko and Vinnikov, 1973, and others) has established the fact that there is a signicant relationship between the precipitation falling on continents and relatively small changes in the mean meridional temperature gradient. Reductions in this gradient that took place, for example, during the warming trend of the 1920's and 1930's, were accompanied by a considerable reduction in precipitation in many intracontinental areas. This led to a substantial increase in the number of droughts covering large areas in the USSR, USA, and other countries, to a drop in the level of many internal water expanses, including the Caspian Sea, and to a number of other consequences.

Change in climatic conditions that can occur by the end of the 20th century as a result of increase in the carbon dioxide concentration will greatly exceed the warming trend of the 1920's and 1930's, and could have a great influence on agricultural productivity. The need for a more detailed study of anthropogenic changes in climate, using models of the dynamics of the atmosphere and oceans, and created just to study possible changes in climate, therefore is obvious. These models should be nonstationary, quite detailed, yet not too cumbersome. Calculations of many variants of change in global climate, decade by decade, then could be made by using them and modern computers. These calculations could take into consideration the influence of the oceans and polar ice, as well as various assumptions as to man's future effect on specific climatic factors. There should be several such models, all incorporating the various assumptions, and coordination of estimates of changes in climate, calculated using them, would enable us to establish the reliability of these estimates and models. Coordination of international scientific efforts in this field is very important.

Calculations using these models will enable us to find the levels, and growth rates, of the anthropogenic effect on global climate that will not exceed the natural variation in the scale of the warming trend of the 1930–1940 period by the end of the 20th century.

There can be no doubt of the great importance of those data for planning the optimum interaction between society and the environment.

## 3. Controlled Effect on Global Climate

One approach this effect can take in order to prevent undesirable changes in climate is based on an artificial change in the concentration of aerosol particles in the lower stratosphere (Budyko, 1974).

The nature of an aerosol in the lower layers of the stratosphere was studied by Junge (1963), Cadle (1972), and others, who established that there is a layer of particles, the mean size of which is measured in tenths of a micron, in the lower stratosphere. The heaviest concentration of particles usually will be observed in a layer about 2 km thick at a height of 18–20 km.

The particles in the stratosphere's aerosol layer are mostly sulfur compounds. Junge hypothesized that these particles are formed from sulfur dioxide ($SO_2$) which, oxidizing during photochemical reactions, is converted into sulfur trioxide ($SO_3$), and with water vapor forms dorplets of surfuric acid ($H_2SO_4$), which comprises the major part of the aerosol in the lower stratosphere (SMIC, 1971).

Aerosol particles are retained in the stratosphere for a long time with the average from 1 to 3 years. The mechanism involved in their precipitation, apparently, is connected with their coagulation and increase in size. The very slow movement of the particles in the vertical direction contrasts with their extremely rapid transfer along the horizontal. Observation of the distribution of particles in the stratosphere after volcanic eruptions shows that when an eruption occurs in the extra-tropical latitudes, the particles are spread over the entire hemisphere within a few months by stratospheric winds. The particles are distributed over both hemispheres within this same time frame if the eruption occurs in the equatorial zone.

According to observational data, aerosol particle concentration can change over broad limits. According to Junge's observations in 1961, for example, the concentration was of the order of magnitude of 0.01 $\mu g/m^3$, and had reached 0.20–0.30 $\mu g/m^3$ in 1968–69. This marked increase in the concentration is explained by the influence of the eruptions of the Agung volcano, and of some others. In 1970, the concentration of aerosol particles decreased to several hundredths of a $\mu g/m^3$.

It is interesting to compare aerosol concentrations for different years with the magnitudes of direct radiation in a cloudless sky for the corresponding periods of time. Z. I. Pivovarova used data from observations made by a group of actinometric stations in the USSR and USA to calculate mean direct radiation magnitudes. These were found to be equal to 98% of the norm in 1961, to 93–94% in 1968–69, and to 97% in 1970. The definite concordance between changes in aerosol concentration and fluctuations in the direct radiation magnitudes is worthy of attention. This makes it possible to take the comparatively few changes in aerosol concentration that are quite characteristic for the whole of the northern hemisphere into consideration.

The direct proportionality between changes in aerosol concentration and changes in radiation flux follows from the theoretical considerations. The data cited in the foregoing then show that a 0.06 $\mu g/m^3$ increase in the aerosol concentration reduces direct radiation by 1%.

This significant dependence of the shortwave radiation flux on the aerosol concentration in the stratosphere explains the marked influence of explosive volcanic eruptions on climatic conditions.

Actinometric observations have shown a marked reduction in direct radiation all over the world, or within the limits of one hemisphere, after explosive eruptions. Direct radiation decreased by 6% after the eruption of Agung volcano in 1963, and there was an even greater reduction in radiation after the eruption of Mount Katmai in 1912, when the reduction in direct radiation was 22% (Budyko, 1971).

Fluctuations in direct radiation after volcanic eruptions are accompanied by changes in summed radiation that make up a comparatively small part of the magnitude of changes in direct radiation. It can be assumed that the influence of a reduction in summed radiation on the temperature of the lower air layer depends on how long the reduction in radiation lasts.

Meteorological observation data show no marked change in the temperature near the earth's surface after the eruption of Agung volcano, but there was a reduction of 0.5°C, on the average, for the northern hemisphere after the eruption of Mount Katmai. A reduction such as this in temperature apparently is a small part of the change that would occur in the event of a prolonged reduction in radiation. These data therefore suggest that the summed radiation decreased by approximately 1% after the first eruption, and by 3% after the second.

The thermal conditions involved in the "warming of the Arctic" were much closer to the steady state. Direct solar radiation increased by approximately 2% in these years, and this increase continued for almost twenty years. An increase such as this in direct radiation corresponds to a 0.3% increase in summed radiation. This should lead to an approximate 0.45°C rise in planetary temperature at the earth's surface under steady-state conditions. This latter value is only slightly greater than the change in temperature established from observational data.

Thus, the change in temperature in the lower air layer when there are fluctuations in stratospheric transmittance depends on two factors: the magnitude of the change in the concentration of aerosol particles in the stratosphere; and the duration of this change.

Let us examine the possibility of a change in climate on the scales observed during the "warming of the Arctic" period, but in the opposite direction; that is, providing for a reduction in the mean temperature at the earth's surface. This would mean a 2% reduction in the mean magnitude of the direct radiation, and this corresponds to a reduction of 0.3% in the summed radiation.

We have already cited an empirical estimate of change in radiation as a function of aerosol concentration in the lower stratosphere. Other calculations of this relationship, based on empirical data and on a theoretical analysis, yield a somewhat greater magnitude of the aerosol mass changing shortwave radiation by the specified magnitude. The mean of all these calculations of the magnitude of the aerosol mass changing summed radiation by 1%, and the direct radiation by 6%, is close to $0.8 \times 10^{-6}$ g cm$^{-2}$.

A close look at this estimate will show that for the purpose indicated above the aerosol mass in the northern hemisphere should be increased by 600,000 tons.

All that has to be done to increase the mass of the particles measured in the lower stratosphere to this magnitude is to deliver 200,000 tons of sulfur to the aerosol layer level, and this, after combustion, will form 400,000 tons of sulfur dioxide. If we consider that the mean life span of aerosol particles in the stratosphere is 2 years, 100,000 tons of sulfur annually would have to be delivered to the lower stratosphere in the northern hemisphere in order to create the necessary concentration of particles consisting of pure sulfuric acid.

The combustion in the stratosphere of much smaller quantities of sulfur than this probably would actually be sufficient to increase the mass of aerosol particles to 600,000 tons, because the hygroscopic droplets of sulfuric acid will absorb water vapor from the air.

The planning of such effects requires accurate estimates of the possible influence on stratospheric aerosols of mass flights by supersonic transport aircraft in the 1980's, because their influence can be considerable (Grobecker, 1973). A study of the causes, and of the nature, of natural fluctuations in the content of these aerosols also is needed for their prediction and consideration when carrying out the effects.

Bringing about the effects on climate examined above will prevent, or weaken, the changes in climate which could take place in 20–30 years as a result of the influence of economic activity. But this effect is not the limit within reach of modern technology. It evidently will be technically possible in the near future to affect the climate, creating changes in it that will be greater by an order of magnitude than those that occurred in the climate in the 1920's–1930's; that is, ensuring a few degree reduction in global temperature. An effect on this scale could be necessary in the 21st century, when the temperature of the lower layers of the atmosphere could rise considerably as a result of a considerable growth in energy production.

The question of the desirability of using effects on climate in the near future, when the influence of economic activity on climate will be comparatively slight, deserves study. Effects that will reduce the mean temperature several tenths of a degree will be economically desirable. Their use to change thermal conditions will have little influence on economic activity, but an increase in the amount of precipitation can be of great practical importance in many areas of insufficient moisture.

Let us note, in conclusion, that elucidation of the possibility, in principle, of a marked change in global climate by comparatively simple means raises a number of questions.

Study of these questions means that attention must be given to the international aspect of the problem of effect on climate. Ye. K. Fedorov (1958, and others) indicated repeatedly that the implementation of any major project affecting climate is possible only on the basis of international collaboration.

The method of affecting climate that we have considered here will lead to change in climatic conditions in the territories of many countries, but the nature of these changes will be different in different regions.

This is what obviously makes it desirable to conclude an international agreement resolving the accomplishment of effect on climate only on the basis of projects that have been examined and approved by responsible international agencies, with the agreement of all interested countries. Such an agreement should encompass measures for guided effect on climate, as well as those types of economic activity similar to mass flights in the lower stratosphere, that could lead to unintentional changes in climatic conditions.

## 4. Global Monitoring of Climatic Factors

A global system of continuous observations and periodic examination of the factors forming global climate, and its anthropogenic changes, is attracting more and more attention from researchers as a component part of global monitoring of the environment. The magnitudes subject to global monitoring, the basic approaches to its accomplishment, approximate accuracies and frequency of measurements, are contained in the well-known SMIC handbook, 1971, and refined in the recently published Global Monitoring Plan project (Munn, 1973). The global monitoring of factors that have an effect on climate is the primary goal of the recently organized network of World Meteorological Organization stations for measuring background atmospheric pollution (Munn, 1973).

Continuous measurement and statistical processing of primary data from natural climate-forming factors are needed for:

1. direct and summed solar radiation, measured on a standard basis by a network of some 1,000 actinometric stations, located in all geographical regions, but unevenly (Budyko, 1971). Most of these stations did not begin operations until the second half of the 20th century, and their data are as yet of little value in determining climatic changes;
2. the albedo, the mean cloud cover, and the mean limit of the polar snow and ice cap. They are most conveniently measured on a global scale with standard weather satellite equipment (television images) (SMIC, 1971). Long series of comparable data must be collected in order to distinguish permanent changes in magnitudes, because these measurements have only been made systematically since the end of the 1960's;
3. the temperature of ocean surfaces, as minimally characterizing thermal inertia and heat transfer in the oceans. SMIC (1971) has pointed out the inadequacy of the study thus far made of the energy exchange between ocean and atmosphere, as well as the need for just a rough working model of that exchange to define more precisely the magnitudes subject to global monitoring. The need for data on the characteristics of the heat exchange that takes place between the surface and deep layers of the ocean and its surface currents is obvious, but it still is impossible to obtain continuous data of this type on a global scale. The temperature of ocean surfaces can be measured from satellites, given certain cloud conditions (SMIC, 1971).

Certain data on factors describing the state of the climate, and on how man has transformed the underlying surface and thus has had an effect on climate, must be obtained periodically, and in definite regions. Included, for example, are glacier mass, ocean level, biomass of forests, area of cities and artificial reservoirs, and the like. Table 7.7 (SMIC, 1971), lists these factors of the "climatic cadastre," the periodicity, and regions to be examined.

The most important of the anthropogenic factors affecting climate are continuous measurements of the carbon dioxide and aerosols in the atmosphere. The $CO_2$ concentration undergoes little latitudinal and seasonal changes, and its global distribution recently was studied in some detail (SMIC, 1971). These changes are greater in the northern hemisphere, where the main seasonally changing sources and flows of atmospheric $CO_2$ into the continental biosphere are located. Therefore, the global increase in the mean $CO_2$ concentration is easy to detect by taking measurements on small islands distant from the continents, or in the southern hemisphere. The main sources of anthropogenic $CO_2$ are in the temperate northern latitudes, while its oceanic flows and sources are in the polar and tropical latitudes, respectively. Measurements of the meridional distribution of the surface concentration of $CO_2$ therefore are necessary by a network of points from the north to the south polar area. The network of WMO base stations for measuring background pollutants in the atmosphere is distributed with this special requirement for station location taken into consideration.

Attempts to estimate the summed annual anthropogenic emission of $CO_2$ into the atmosphere, and its increase in recent decades, were published recently (Keeling, 1973). A comparison between these estimates and the measured increase in the content of $CO_2$ in the atmosphere possibly will permit refining the rate at which absorption of $CO_2$ by the ocean takes place, as well as how it has changed over the years. The importance of monitoring this value is obvious, as is how it changes and depends on the state of the ocean surface. We must, therefore, make systematic measurements of the $CO_2$ content in the surface waters of the oceans in different latitudes simultaneously with its content in the air above the water, in order to estimate the rate of $CO_2$ exchange between ocean and atmosphere, and to study its change. The necessary frequency of these measurements in time and space still remains to be determined. The establishment of the physicochemical causes of its change will be the primary goal of these measurements and studies of the rate of $CO_2$ exchange.

Global monitoring of atmospheric aerosols is more complicated. The measurements of atmospheric turbidity recommended and made by the network of WMO stations are insufficient for distinguishing the content of stratospheric and tropospheric aerosols, the effect of which on climate, as indicated above, differs somewhat. Individual, systematic (at least once a season) measurements of the content of aerosols in the stratosphere up to the 30 km level, using aircraft and balloons, with determination of the chemical composition of the aerosols, of the type of programs now conducted in the USA therefore are needed (Cadle, 1972). These measurements are particularly important after major volcanic eruptions. They also will be urgently needed to establish the possibility of global regulation of the climate by the method proposed in paragraph 3.

It is very difficult to establish permanent changes in mean values on a global scale because of the great changeability and diversity of the aerosol concentrations in the troposphere, particularly over areas of intensive economic activity. Time-consuming processing of a great volume of data from a dense network of stations is required in order to obtain reliable statistical characteristics. Measurements of atmospheric turbidity, and analyses of filtered aerosol samples obtained by the network of WMO stations in terms of background atmospheric pollution certainly will be inadequate for this purpose.

Ground measurements of mean global levels of the content of aerosols in the troposphere can be made most conveniently in the polar region, for from the main sources of anthropogenic and natural aerosols. Attempts are being made to estimate the global content of aerosols in the atmosphere from measurements in Antarctica and Greenland, as well as from those made by research ships on the high seas.

Data on the intensity and special features of global transfer in different parts of the atmosphere are needed for these estimates, and can be obtained by using tracers; radioactive, or those like the inert gas freon-11 (a halogenated hydrocarbon, $CCl_3F$) of anthropogenic origin, recently measured by a research ship in the oceans of the southern hemisphere (Lovelock, 1973).

Study and monitoring of the global, more or less homogeneous, layer of aerosols in the middle, and particularly in the upper, troposphere, and which are primarily of continental origin, also can be done by aircraft sounding in the meridional plane. The lower level of seasonal aircraft measurements of the content of stratospheric aerosols indicated above can be omitted for this purpose.

Finally (and this is very important), detailed climatological processing and analysis of standard meteorological observations (temperature and precipitation in particular) is necessary to distinguish climatic fluctuations now, and in the near future, as these investigations are continued for the end of the last century, and the first half of this one (Rubinshteyn, Ye. S., 1973).

All magnitudes measured should be associated as much as possible with the quantitative models of global climate discussed above. These magnitudes, and the frequency and accuracy with which they are measured, should meet the requirements for initial data, and should be capable of being compared with the resultant calculated climatic parameters of the models.

## BIBLIOGRAPHY

Budyko, Mi. I., 1969, *Izmeneniya klimata* [Climatic Changes], Gidrometizdat, Leningrad, 35 p.

Budyko, M. I., 1971, *Klimat i zhizn'* [Climate and Life], Gidrometeoizdat, Leningrad, 470 p.

Budyko, M. I., 1972, *Vliyaniye cheloveka na klimat* [Man's Influence on Climate], Gidrometeoizdat, Leningrad, 47 p.

Budyko, M. I., 1973, *Atmosfernaya uglekislota i klimat* [Atmospheric Carbon Dioxide and Climate], Gidrometeoizdat, Leningrad, 32 p.

Budyko, M. I., 1974, "An Approach to Affecting Global Climate," *Meteorologiya i gidrologiya*, No. 2.

Budyko, M. I., Vasishcheva, M. A., 1971, "The Influence of Astronomical Factors on Quarternary Glaciation," *Meteorologiya i gidrologiya*, No. 6, pp. 37–47.

Budyko, M. I., Vinnikov, K. Ya., 1973, "Modern Climatic Changes," *Meteorologiya i gidrologiya*, No. 9, pp. 3–13.

Kondrat'yev, K. Ya., et al., 1973, *Vliyaniye aerozolya na perenos izlucheniya: vozmozhnyye klimaticheskiye vozdeystviya* [The Influence of Aerosols on Radiation Transfer: Possible Climatic Effects] Izd. LGU, 266 p.

Rubinshteyn, Ye. S., 1973, *Struktura kolebaniy temperatury na severnom polusharii* [The Structure of Temperature Fluctuations in the Northern Hemisphere], Gidrometeoizdat, Leningrad, 34 p.

Federov, Ye. K., 1958, "The Effect of Man on Meteorological Processes," *Voprosy filosofii*, No. 4.

Berger, A. L. 1973. Theorie astronomique des paleoclimats, v. I, v.II, Louvain.

Cadle, R. D., 1972. Composition of the stratospheric "sulphate layer," EOS. Trans. Amer. Geoph. Union, v. 55, No. 9, 812.

Grobecker, A. J., 1973. United States Department of Transportation. Research program for assessment of stratospheric pollution. DOT TST-2.1, Wash., D.C. 22 p.

Junge, C. E., 1963. Air chemistry and radioactivity. Acad. Press.

Keeling, C. H., 1973. Industrial production of carbon dioxide from fossil fuels and limestone. Tellus. v. 25, No. 2, p. 174–198.

Krey, P. et al., 1973. Project Airstream. Fallout program. Quart. summary report. HASL-276, p. II-6–II-70.

Lovelock, J. E. et al., 1973. Halogenated hydrocarbons in and over the Atlantic. Nature, v. 241, No. 5386, p. 194–196.

Machta, L., 1973. Man's influence on the climate. A status report. Lecture at the VI Session of Commiss. Atm. Sci. WMO.

Munn, R. E., 1973, Global environmental monitoring system. Action Plan for Phase I. SCOPE Rep. 3, Toronto.

Reiter, E. R., 1971. Atmospheric transport processes. P. 2. Chemical tracers. U.S. A.E.C. TID 25314, Oak Ridge, 382 p.

SMIC, 1971. Inadvertent climate modification. MIT Press, Cambr., Mass. and Lond. 308 p.

Schneider, S. H. & Gal-Chen, T., 1973. Numerical experiments in climate stability. Journ. Geoph. Res., v. 78, No. 27, p. 6182–94.

Sellers, W. D., 1969. A global climate model based on the energy balance of the earth-atmosphere system. J. Appl. Met., v. 8, No. 3, p. 392–400.

Smagorinsky, J., 1974. Global atmospheric modeling and the numerical simulation of climate. In Weather Modification. Ed. W. N. Hess, Wiley & S, N.Y.

Washington, W. M., 1971. On the possible use of global atmospheric models for the study of air and thermal pollution. In Man's Impact on the Climate. Cambridge MIT Press p. 265–276.

Yamamoto, G. & Tanaka, M., 1972. Increase of global albedo due to air pollution. J. Atm. Sci., v. 29, No. 8, p. 1405–12.

# Concluding Remarks for the U.S.-U.S.S.R. Symposium on Comprehensive Analysis of the Environment

Leland D. Attaway *

Comprehensive Analysis of the Environment involves five major problem areas:

- Pollutant Characterization: the identification of problem pollutants and the description of their sources, emission characteristics, and transformation processes.
- Route and Fate: description of the relationship between pollution emissions and ambient environmental quality (including intermediate and ultimate biological locations).
- Effects: the relationship between pollution concentrations at the receptor (victim) and its effect upon that receptor, both health and ecological. Indirect effects upon economic and social systems must be considered.
- Monitoring and Pollution Reduction Methods: instrumentation for assessing emissions and ambient quality, and technological and operational means for reducing emissions.
- Policy: the setting of emission or ambient pollutant standards, land use prescriptions, etc., that is, the goals of comprehensive environmental analysis.

These problems run the range of detail from micro (e.g., human cell processes) through macro (e.g., land use allocations in a large region). We are essentially at the beginning of developing knowledge, tools and data in all of them. Therefore, one conclusion we can safely draw is: we need advances, and international cooperation, at all levels in all these problem categories.

Dr. Izrael has introduced in his paper on Complex Analysis of the Environment a comprehensive mathematical statement of the environmental analysis problem; a mathematical formula which encompasses all the above components, at all levels of aggregation from micro to macro. It has two important potential uses:

(1) Organizational: to identify all essential components of the problem; to indicate their interrelationships, so that research products on individual components are designed to fit one another overall; to determine the research priorities, by identifying the strongest and weakest links in the knowledge chain; and to coordinate the overall research program.

(2) Analytic: to actually perform special analyses. However, in its present form the mathematical statement is more representative than analytical; the form taken for applications will depend upon the specific objectives and the state of knowledge at the time.

We must use some such statement for both purposes—but we must be making applications immediately, and so cannot wait until our knowledge permits us to use such a comprehensive expression. Indeed, it is clear that the equation as it stands will never be used in its analytic role—it simultaneously has too much scope and too much detail. One or the other has to be reduced, depending upon the specific application.

The regional analysis of Lake Baykal, discussed in the paper System Analysis and Imitation Mathematical Modelling, by Dr. Izrael et al., represents what can be done today in Comprehensive Environmental Analysis for a region: a group of sector models (for example: air and water transport; land use; pollution reduction methods) which can be applied independently of one another, but which feed one another inputs. They include all the components I listed earlier: pollutant characterization; pollutant route and fate; effects; monitoring and pollution reduction methods; and policy, such as standard setting and land use allocations.

The product of applying these tools for a specific region must be a comprehensive plan for that region covering the next ten or fifteen years. The environment is but one aspect of that plan. It should also include guidance on:

(1) Environmental goals (expressed as ambient air quality standards, water standards, land-use guidelines, etc.)

(2) Effluent standards, such as flue gas scrubber control guidelines.

(3) Economic growth goals.

(4) Land use.

---

*Deputy Assistant Administrator for Program Integration, U.S. Environmental Protection Agency

(5) Transportation.

(6) Utilities (water supply, power, waste collection and treatment).

(7) Population growth and distribution.

All these other (than environmental) aspects must be such as to meet the environmental goals over the ten to fifteen year period. However, if the cost-benefit balance point so indicates, then the environmental goals may have to be relaxed; or the population goals; or the economic goals. This plan must be flexible enough to change annually, and it must be a pragmatic blend of proceeding "today" with the tools and knowledge available. It will be a blend of *Best Practical Technology* application and *ambient* proscriptions.

It is not feasible to set pollutant standards independently of cost-benefit considerations. All pollution cannot be terminated, so clearly some pollution is worth the attendant social costs. The lowest level of visible organism response will often be too costly a level for a standard, given the undamaging nature of the effect. The standards sometimes must be set at more relaxed levels closer to, but still safely below, harmful effect levels—at least as an interim measure. We thus must consider "standards" which change over calendar time, in order to move closer and closer to our ultimate goals but at a pace which will not unacceptably damage the economy. But we *must converge* to those ultimate goals.

Within the context of my earlier (just above) remarks, we would suggest the following candidates as high priority joint projects. In all cases, maximum use should be made of the product from Working Groups already established within the rest of the Environmental Agreement. Indeed, the principal role of this Comprehensive Environmental Analysis Project should be stimulation of integrated working group efforts. The candidates are:

(1) Pathways, fates and effects of multi-media pollutants, and related multi-media standards.

(2) Requirements for multiple standards regulating long and short term doses.

(3) Responses of aquatic and terrestrial ecosystems to pollutants.

(4) Analysis techniques and measurement methods for biological and genetic effects.

(5) Techniques for evaluating socio-economic costs and benefits.

(6) Evaluation of worth and feasibility of a joint system for assessing pollution concentrations in the biosphere.

☆ U.S. GOVERNMENT PRINTING OFFICE : 1976 O—587-798